

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL: U09269; ABA9457.1; -
DR EMBL: U74683; AAB58400.3; -
DR HSSP: O46427; 8PCH.
DR MEROPS: C01.070; -.
DR MCD: MGI:109553; Clsc.
DR InterPro: IPR000668; Peptidase_C1.
DR InterPro: IPR000169; Thiolprot_act_site.
DR Pfam: PF00112; Peptidase_C1; 1.
DR PRINTS: PR00705; PAPAIN.
DR PROSITE: PS00139; THIOL_PROTEASE_CYS; 1.
DR PROSITE: PS00639; THIOL_PROTEASE_HIS; 1.
DR PROSITE: PS00640; THIOL_PROTEASE_ASN; 1.
KW Hydrolyase; Thiol protease; Lysosome; Glycoprotein; Zymogen; Signal.
FT SIGNAL 1 24
FT PROPEP 25 230
FT CHAIN 231 393
FT CHAIN 394 462
FT ACT_SITE 257 257
FT ACT_SITE 404 404
FT ACT_SITE 426 426
FT CARBOHYD 29 29
FT CARBOHYD 25 25
FT CARBOHYD 275 275
SQ SEQUENCE 462 AA; 52376 MW; 56574B38D7DF4710 CRC64;

Query Match 7.1%; Score 89; DB 1; Length 462;
Best Local Similarity 23.6%; Pred. No. 1.1;
Matches 35; Conservative 19; Mismatches 52; Indels 42; Gaps 7;

OY 56 ACPELTCEDEPENCIPGSIIVRYMLANFALANDNGWGNPPLKHFDAEPALQIA---- 111
DB 319 SCFPYTAKDSP--CKPRENCLRYSSDYVVGCGGNCALMKLELVKRPANVAFEVH 376
OY 112 ----QYAGIVPAVFRVPCPEKGGGIFRTINGNDY--FDLVLTNNGAGDIRAVSLKGS 165
DB 377 DDFLHYHSGI-----YHHTGLSDPFNFEL---TN-----HAVLLVGY 411
OY 166 KTD-----QWOSMRNGOMNOMNTYL 188
DB 412 GRDPTVGTIEWIIKNSWMSWGESGIFR 439

RESULT 11
GUN_MYTED
ID GUN_MYTED STANDARD: PRT: 181 AA.
AC P82186;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Endoglucanase (EC 3.2.1.4) (Endo-1,4-beta-glucanase) (Cellulase)
DE (CMCAB).
OS Mytilus edulis (Blue mussel).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
OC Mytiloidea; Mytilidae; Mytilus.
OX NCBI_TaxID=6550;
RN [1]
RP SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.
RC TISSUE-Digestive gland;
RX MEDLINE=20389517; PubMed=10931178;
RA Xu B., Hellman U., Ersson B., Janson J.-C.;
RT Purification, characterization and amino-acid sequence analysis of a
RT thermostable, low molecular mass endo-beta-1,4-glucanase from blue
RL Eur. J. Biochem. 267:4970-4977(1999).

```

```

CC -----
CC FUNCTION: ACTIVE TOWARDS THE SOLUBLE CARBOXYMETHYLCELLULOSE (CMC).
CC POSSIBLES EXPANSIN ACTIVITY TOO.
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
CC linkages in cellulose.
CC -1- TISSUE SPECIFICITY: DIGESTIVE GLAND.
CC -1- MASS SPECTROMETRY: MW=19702; METHOD=MALDI.
CC -1- MISCELLANEOUS: Has an isoelectric point of 7.6. Its optimum pH is
CC 4.6 and optimum temperature is between 30-50 degrees Celsius.
CC -1- SIMILARITY: BELONGS TO CELLULOSE FAMILY K (FAMILY 45 OF GLYCOSYL
CC HYDROLASES).
DR InterPro: IPR000334; Glyco_hydro_45.
DR PROSITE: PS01140; GLYCOSYL_HYDROL_F45; FALSE_NEG.
KW Cellulose degradation; Hydrolase; Glycosidase.
FT ACT_SITE 24 24
FT ACT_SITE 132 132
FT DISULFID 4 16
FT DISULFID 30 69
FT DISULFID 32 176
FT DISULFID 65 178
FT DISULFID 72 157
FT DISULFID 103 113
SQ SEQUENCE 181 AA; 19711 MW; E00A8C57203823F6 CRC64;

Query Match 7.0%; Score 87.5; DB 1; Length 181;
Best Local Similarity 25.2%; Pred. No. 0.52;
Matches 36; Conservative 9; Mismatches 43; Indels 55; Gaps 7;

OY 8 GHATFYGADASCTMG-----GACGYGNLHSGY-----GLDTAALSTALPNSGCK---- 53
DB 6 GNPRRYNGSCASTYVYHSHKACGCGPASGDAQCGMNAAGSVVAASQMYTDSGKKGWC 65
OY 54 ----CGAGFELT-----CED---DPEWCTPGSIIVRYMLA 81
DB 66 GCHCGGCIKLTGGYVPGGCGPVREGLSKTEWITMLCPNIPVNDQWCMGSGQYGGHMKY 125
OY 82 NFALND--NG-----GCMNP 95
DB 126 GYELHLDLENGRQGYVMGMNMP 148

RESULT 12
MPL2_LOLPR
ID MPL2_LOLPR STANDARD: PRT: 97 AA.
AC P14947;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Pollen allergen Lol p 2-A (Lol p II-A).
DE Pollen perenne (Perennial ryegrass).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Poaceae; Lolium.
OX NCBI_TaxID=4522;
RN [1]
RP SEQUENCE.
RX MEDLINE=89291864; PubMed=2472390;
RA Ansari A.A., Shenbagamurthi P., Marsh D.G.;
RT "Complete amino acid sequence of a Lolium perenne (perennial rye
RT grass) pollen allergen, Lol p II."
RL J. Biol. Chem. 264:11181-11185(1989).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DISEASE: CAUSES GRASS POLLEN ALLERGY.
CC -1- SIMILARITY: BELONGS TO THE LOI p I FAMILY OF ALLERGENS.
CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE CBD DOMAIN.
DR PIR: A34291; A34291.
DR HSSP: P43214; IMHO.
DR InterPro: IPR000882; Pollen_allergen.
DR Pfam: PF01357; Pollen_allergen; 1.
DR ProDom: PD002179; Pollen_allergen; 1.
DR PROSITE: PS00843; EXPANSIN_CBD; 1.
KW Allergen; Multigene family.
FT DOMAIN 15 96
FT VARIANT 29 29 S->A.

```


ID CATC_HUMAN STANDARD: PRT: 463 AA.
AC P53634;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Dipeptidyl-peptidase I precursor (EC 3.4.14.1) (DPP-I) (DPP1)
DE (Cathepsin C) (Cathepsin J) (Dipeptidyl transferase).
GN CTSC OR CPPI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_Taxid=9606;
RN [1]
RN SEQUENCE FROM N.A.
RC TISSUE=ileum;
RX MEDLINE=95377428; PubMed=7649281;
RA Paris A., Struckel B., Pungertar J., Renko M., Dolenc I., Turk V.;
RT "Molecular cloning and sequence analysis of human preprocathepsin C";
RL FEBS Lett. 369:326-330(1995).
RP SEQUENCE FROM N.A.
RX MEDLINE=97248590; PubMed=9092576;
RA Rao N.V., Rao G.V., Hoidal J.R.;
RT "Human dipeptidyl-peptidase I. Gene characterization, localization,
RT and expression";
RL J. Biol. Chem. 272:10260-10265(1997).
RN [3]
RP VARIANTS PLS F-249; L-252; P-272; S-301; C-339 AND C-347.
RX MEDLINE=20047769; PubMed=10581027;
RA Toomes C., James J., Wood A.J., Wu C.L., McCormick D., Lench N.,
RA Hewitt C., Moynihan L., Roberts B., Woods C.G., Markham A., Wong M.,
RA Widmer R., Gaftar K.A., Pemberton M., Hussein I.R., Temtamy S.A.,
RA Davies R., Read A.P., Sloan P., Dixon M.J., Thacker N.S.;
RT "Loss-of-function mutations in the cathepsin C gene result in
RT peridontal disease and palmoplantar keratosis";
RL Nat. Genet. 23:421-424(1999).
RN [1]
RN FUNCTION: THIOL PROTEASE. HAS DIPEPTIDYLPEPTIDASE ACTIVITY. CAN
RN DEGRADATE GLUCAGON.
RN [1]
RN CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Xbb-|-
RN xcc, except when xaa is Arg or Lys, or xbb or xcc is Pro.
RN [1]
RN COFACTOR: REQUIRES CHLORIDE IONS FOR ACTIVITY.
RN [1]
RN SUBUNIT: DIMER OF AN ALPHA CHAIN AND A BETA CHAIN CROSS-LINKED
RN BY A DISULFIDE BOND.
RN [1]
RN SUBCELLULAR LOCATION: Lysosomal.
RN [1]
RN DISEASE: DEFECTS IN CTSC ARE A CAUSE OF PAPILLON-LEFEVRE SYNDROME
RN (PLS) ALSO KNOWN AS KERATOSIS PALMOPLANTARIS WITH
RN PERIODONTOPATHIA. IT IS AN AUTOSOMAL RECESSIVE DISORDER THAT IS
RN MAINLY ASCERTAINED BY DENTISTS BECAUSE OF THE SEVERE PERIODONTITIS
RN THAT AFFECTS PATIENTS. BOTH THE DECIDUOUS AND PERMANENT
RN DENTITIONS ARE AFFECTED, RESULTING IN PREMATURE TOOTH LOSS.
RN PALMOPLANTAR KERATOSIS, VARYING FROM MILD PSORIASIFORM SCALY SKIN
RN TO OVERT HYPERKERATOSIS, TYPICALLY DEVELOPS WITHIN THE FIRST THREE
RN YEARS OF LIFE. KERATOSIS ALSO AFFECTS OTHER SITES SUCH AS ELBOWS
RN AND KNEES.
RN [1]
RN SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1; ALSO KNOWN AS THE
RN PAIN FAMILY OF THIOL PROTEASES.
RN [1]
RN This Swiss-Prot entry is copyright. It is produced through a collaboration
RN between the Swiss Institute of Bioinformatics and the EMBL outstation -
RN the European Bioinformatics Institute. There are no restrictions on its
RN use by non-profit institutions as long as its content is in no way
RN modified and this statement is not removed. Usage by and for commercial
RN entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
RN or send an email to license@isb-sib.ch).
RN [1]
RN EMBL: X87212; CA60671.1; -;
RN EMBL: U79415; AAC51341.1; -;
RN HSSP: O46427; 8PCH.
RN MEROPS: C01.070; -;
RN MIM: 602365; -;
RN MIM: 245000; -;
RN InterPro: IPR000668; Peptidase_C1.
RN InterPro: IPR000169; Thiolprol_act_site.

DR Pfam: PF00112; Peptidase_C1. 1.
DR PRINTS: PR00705; PAPA1N.
DR PROSITE: PS00139; THIOL_PROTEASE_CYS. 1.
DR PROSITE: PS00639; THIOL_PROTEASE_HIS. 1.
DR PROSITE: PS00640; THIOL_PROTEASE_ASN. 1.
DR Hydroxylase: Thiol protease; Lysosome; Glycoprotein; Zymogen; Signal;
KW Disease mutation.
FT SIGNAL 1 24
FT PROPEP 25 230
FT CHAIN 231 394
FT CHAIN 395 463
FT ACT_SITE 258 463
FT ACT_SITE 405 405
FT ACT_SITE 427 427
FT CARBOHYD 29 29
FT CARBOHYD 53 53
FT CARBOHYD 119 119
FT CARBOHYD 216 216
FT CARBOHYD 249 249
FT VARIANT 252 252
FT VARIANT 272 272
FT VARIANT 301 301
FT VARIANT 339 339
FT VARIANT 347 347
FT VARIANT 347 347
FT VARIANT 51841 MM; 759BSEF1290C3771 CRC64;
SQ SEQUENCE 463 AA; 51841 MM; 759BSEF1290C3771 CRC64;
Query Match 6.7%; Score 83; DB 1; Length 463;
Best Local Similarity 25.3%; Pred. No. 3.6;
Matches 37; Conservative 9; Mismatches 62; Indels 36; Gaps 6;
QY 56 ACFELTCEDDPEMKIPISIVRYNLANFALANDNGMCNPLKHFDLAEPALQIA---- 111
DB 320 ACFPYTGDSF--CKMKEDCFRYSSEHYVGGFYGCNEALMKLELVHNGPMAVFEVY 377
QY 112 ---QYRAGIVPAFRVRVPCCKGGIRFTINGNPFDYLVITVGAGAGDIRAVSLKGSY 167
DB 378 DDFLHYKKGIYHNTGLRDPF-----NPF-----ELTN-----HALLVGYGT 414
QY 168 DQNGSM-----SRWQGNQMSNTLYLR 188
DB 415 DSASGMVYIVKXSWGTGCGENGFR 440
RESULT 15
GN2_TRIPE STANDARD: PRT: 418 AA.
AC P07962;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Endoglucanase EG-II precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase)
DE (Cellulase)
GN BGL2 OR BGLII.
OS Trichoderma reesei (Hypocrea jecorina).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreales; Hypocreaceae; Hypocrea.
OX NCBI_Taxid=51453;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN-VTT-D-80133;
RX MEDLINE=88255850; PubMed=3384334;
RA Saloheimo M., Lehtovaara P., Penttilae M., Teeri T.T., Staahlberg J.,
RA Johansson G., Pettersson G., Clavenssens M., Tonne P., Knowles J.K.C.;
RT "EgIII, a new endoglucanase from Trichoderma reesei: The
RT characterization of both gene and enzyme";
RL Gene 63:11-21(1988).
RN [2]

RP ACTIVE SITE GLU-350.
 RX MEDLINE=93131031; PubMed=8093602;
 RA Macarron R., van Beunen J., Henriessat B., de la Mata I.,
 RA Claessens M.,
 RT "Identification of an essential glutamate residue in the active site
 RT of endoglucanase III from *Trichoderma reesei*.";
 RL FEBS Lett. 316:137-140(1993).
 CC -1- FUNCTION: THE BIOLOGICAL CONVERSION OF CELLULOSE TO GLUCOSE
 CC GENERALLY REQUIRES THREE TYPES OF HYDROLYTIC ENZYMES:
 CC (1) ENDOGLUCANASES WHICH CUT INTERNAL BETA-1,4-GLUCOSIDIC BONDS;
 CC (2) EXOCELLULOBIOTRYPHANOLASES THAT CUT THE DISSACCHARIDE CELLULOSE
 CC FROM THE NONREDUCING END OF THE CELLULOSE POLYMER CHAIN;
 CC (3) BETA-1,4-GLUCOSIDASES WHICH HYDROLYZE THE CELLULOSE AND OTHER
 CC SHORT CELLO-OLIGOSACCHARIDES TO GLUCOSE.
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
 CC linkages in cellulose.
 CC -1- SIMILARITY: BELONGS TO CELLULOSE FAMILY A (FAMILY 5 OF GLYCOSYL
 CC HYDROLASES).
 CC -1- SIMILARITY: CONTAINS 1 FUNGAL-TYPE CELLULOSE-BINDING DOMAIN (CBD).
 CC -1- CAUTION: WAS ORIGINALLY CALLED ENDOGLUCANASE EG-III.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M19373; AAA34213.1; -
 CC PIR: S28372; S28372.
 DR HSP: P00725; 2CBH.
 DR InterPro: IPR000254; CBD_fungal.
 DR InterPro: IPR001547; Glyco_hydro_F5.
 DR Pfam: PF00734; CBD_1; 1.
 DR Pfam: PF00150; cellulase; 1.
 DR ProDom: PD001821; CBD_fungal; 1.
 DR SMART: SM00236; fcbd; 1.
 DR PROSITE: PS00562; CBD_FUNGAL; 1.
 DR PROSITE: PS00659; GLYCOSYL_HYDROL_F5; 1.
 DR Cellulose degradation; Hydrolyase; glycoprotein; signal.
 KW CELLULOSE
 FT CHAIN 1 21
 FT SIGNAL 22 418
 FT DOMAIN 22 57 ENDOGLUCANASE EG-II.
 FT DOMAIN 58 91 CELLULOSE-BINDING (BY SIMILARITY).
 FT DOMAIN 92 418 LINKER.
 FT MOD_RES 22 22 CATALYTIC.
 FT CARBOHYD 124 124 PYRROLIDONE CARBOXYLIC ACID.
 FT DISULFID 29 46 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT ACT_SITE 40 56 BY SIMILARITY.
 FT ACT_SITE 239 239 PROTON DONOR (BY SIMILARITY).
 FT ACT_SITE 350 350 NUCLEOPHILE.
 SQ SEQUENCE 418 AA; 44227 MW; 26A492D55237A49B CRC64;

Query Match 6.5%; Score 81; DB 1; Length 418;
 Best local Similarity 19.9%; Pred. No. 4.9; 77; Indels 152; Gaps 15;
 Matches 64; Conservative 28; Mismatches 77;

QY 13 YGADADAGTGMGACGYGNLHSGYGLDTALSTALFNSGKC--GACFELCEDDEPEW- 68
 DB 16 YGAVAAOQTYWGGCG-----GIG-----WSGPTNCAPGSACTL-----NPTVA 54
 QY 69 -CIPGSI-----VRY-----NLANFALANDNGWC- 93
 DB 55 OCIFGATITITSTRPSPGTTTTRATSTSSPTPTSSGVAFAGVNTAGFDGCTTDTGTCV 114
 QY 94 ---NPLKHFDAEPALDIAOTRAGI-----VPAFRRVCEKEGGI 133
 DB 115 TSKVYPLKNTGSGN-----NYPDGIQOMHFVNEDGMTIRFLPVGWQYLVNNGN 167
 QY 134 RFTINGNPFYDLVL-----TTNAGAGDIRAVSLKSGKTDQWQ 171
 DB 168 LNSTSTSKYDQLVGGCLSLAYCIVDIHNARANGGIIGGPTPMAQFTSLWSQLASKYA 227

QY 172 SMSRW-----GQNNQS-NTYLRGOS 191
 DB 228 SOSRWEGIMNEPMDVNTNMTATVGEVYTAIINAGATSGFISLPENDWQSAAGAFISDS 287
 QY 192 LS--FOYVTSDDRTV-VSYDV 209
 DB 288 AAALSQVTNPDGSTTNLITFDV 308

Search completed: October 11, 2002, 14:57:33
 Job time: 6.81181 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 11, 2002, 14:53:05 : Search time 11.2915 Seconds

(without alignments)
1914.721 Million cell updates/sec

Title: US-09-896-301-5

Perfect score: 1245

Sequence: 1 DNGCMERGHATFYCGADASG.....SYDVVPHDMQRCGFEEGGQF 225

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_71.*

1: PIR1.*

2: PIR2.*

3: PIR3.*

4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1203.5	96.7	255	2	T50656 expansin EXP2 [imp
2	962	77.3	250	2	T110079 expansin S1 precu
3	958.5	77.0	253	2	F84831 probable expansin
4	927	74.5	232	2	T09821 expansin (clone pp
5	927	74.5	232	2	T09826 expansin (clone pp
6	923	74.1	232	2	T09818 expansin (clone pp
7	912	73.3	232	2	T09825 expansin - upland
8	912	73.3	232	2	T09786 expansin - rice
9	884	71.0	246	2	T04175 expansin 2 - rice
10	883	70.9	251	2	T03298 expansin EXP1 [imp
11	869	69.8	237	2	T50654 probable expansin
12	868	69.7	248	2	D84444 probable expansin
13	851	68.4	257	2	D84820 probable expansin
14	850	68.3	252	2	T02530 probable expansin
15	846	67.8	260	2	T47689 pollen allergen ho
16	844	67.0	258	2	S53082 expansin 9 [imp
17	834	66.9	257	2	T02727 probable expansin
18	833	66.9	257	2	T50658 expansin EXP6 [imp
19	832	66.8	259	2	T50653 expansin 18 - toma
20	832	66.8	255	2	T06573 expansin 1 - toma
21	828	66.5	261	2	T07630 alpha-expansin OSE
22	825	66.3	264	2	T50659 alpha-expansin OSE
23	825	66.3	258	2	T48247 alpha-expansin OSE
24	819	65.8	260	2	T08016 probable expansin
25	819	65.8	262	2	T50660 alpha-expansin 2 [
26	814	65.0	258	2	T10083 expansin S2 precu
27	809.5	65.0	255	2	T50655 expansin EXP5 [imp
28	761	61.1	255	2	T02010 expansin homolog T
29	759	61.0	252	2	F86335 hypothetical prote

30	758.5	60.9	261	2	T03737 expansin - rice
31	694.5	55.8	255	2	T03299 expansin 3 - rice
32	619	49.7	160	2	T09871 expansin - upland
33	609	48.9	257	2	G96654 hypothetical prote
34	602	48.4	257	2	F86259 protein T12C24.10
35	532	42.7	256	2	expansin homolog F
36	287	23.1	77	2	T09815 expansin (clone pb
37	278.5	22.4	262	2	S38620 allergen Phl p 1 -
38	272	21.8	102	2	T09828 allergen Phl p 1 -
39	272	21.8	263	2	S44182 beta-expansin - ri
40	266.5	21.4	491	2	T04301 beta-expansin - soy
41	266	21.4	261	2	T04301 beta-expansin [imp
42	261	21.0	277	2	S48032 beta-expansin [imp
43	259	20.8	259	2	T50657 beta-expansin [imp
44	259	20.8	271	2	H84592 beta-expansin [imp
45	257	20.6	263	2	S13614 major allergen Lol

ALIGNMENTS

RESULT 1

T50656 expansin EXP2 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 28-Jul-2000

C:Accession: T50656

R:Shcherban, T.Y.; Shi, J.; Duracko, D.M.; Gullinan, M.J.; McQueen-Mason, S.J.; Shi

Proc. Natl. Acad. Sci. U.S.A. 92, 9245-9249, 1995

A:Title: Molecular cloning and sequence analysis of expansins--a highly conserved, mu

A:Accession: T50656

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-255 <SHC>

A:Cross-references: EMBL:U30481; PDB:1AMB38073.1

A:Experimental source: variety Columbia

C:Genetics:

A:Gene: EXP2

C:Function:

A:Description: Induces extension (creep) in plant cell walls

C:Superfamily: expansin

C:Keywords: cell wall

Query Match

Best Local Similarity

Matches 220: Conservative 1; Mismatches 4; Indels 3; Gaps 1;

Score 1203.5; DB 2; Length 255;

Pred. No. 9.1e-102;

1 DNGCMERGHATFYCGADASGTMGACGYGNLSQGYGLQTAALSTALFNSGKCGACFEL 60

28 DNGCMERGHATFYCGADASGTMGACGYGNLSQGYGLQTAALSTALFNSGKCGACFEL 87

61 TCDDDEKICPGSIIV---RYNLANFALANDNGCMNPLKFDLAEPFLQIAQYRACI 117

88 TCDDDEKICPGSIIVSATNFCPPNFALANDNGCMNPLKFDLAEPFLQIAQYRACI 111

118 VPAFAFRVPCCEKGGIRFTINCPYFDVLITNVGAGGIRAVLSKSTTDWQSMNRW 177

148 VPAFAFRVPCCEKGGIRFTINCPYFDVLITNVGAGGIRAVLSKSTTDWQSMNRW 207

178 GNMOSNTYLRGQSLFQVTDSDGRTVVSVDVPHDMQRCGFEEGGQF 225

208 GNMOSNTYLRGQSLFQVTDSDGRTVVSVDVPHDMQRCGFEEGGQF 255

RESULT 2

T10079 expansin S1 precursor - cucumber

C:Species: Cucumis sativus (cucumber)

C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000

C:Accession: T10079

R:Shcherban, T.Y.; Shi, J.; Duracko, D.M.; Gullinan, M.J.; McQueen-Mason, S.J.; Shi

Proc. Natl. Acad. Sci. U.S.A. 92, 9245-9249, 1995

A:Title: Molecular cloning and sequence analysis of expansins--a highly conserved, multi
 A:Reference number: Z14894; MUID:96016146
 A:Accession: J10079
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-250 <SHC>
 A:Cross-references: EMBL:U30382; NID:g1040874; PIDN:AAB37746.1; PID:g1040875
 A:Experimental source: cultivar Burpee Pickler
 A:Genetics:
 A:Gene: Exp1
 A:Function:
 A:Description: mediates cell wall extension
 C:Superfamily: expansin
 F:1-23/Domain: signal sequence #status predicted <SIG>
 F:24-250/Product: expansin #status predicted <MAT>

Query Match 77.3%; Score 962; DB 2; Length 250;
 Best Local Similarity 75.0%; Pred. No. 7,4e-80;
 Matches 171; Conservative 26; Mismatches 27; Indels 4; Gaps 2;

QY 1 DNGGWHATFFYGADAGTGMGACGYGNLHSGYGLQTAALSTALFNSGQKGCACFEL 60
 DB 24 DYGWGHATFFYGADAGTGMGACGYGNLYSGYGTNTVALSTALFNNGLSCGACFEM 83

QY 61 TCEDDEPWCIPGSIIV---RYNLANFALANDNGMCNPLKHPDLAPAFLOIAQYRAGI 117
 DB 84 TCTNDPCKMCLPCTIRYATATNFCPPNFALPNNNGMCNPLDHPMAEPALQIAQYRAGI 143

QY 118 VVAFAFRVCEKGGGIRFTINGNPFEDLVLTITNGAGADIRAVSLKSKTDQMSMRNW 177
 DB 144 VVPSFRVCEKMGKGGIRFTINGNPFEDLVLTITNGAGADIRAVSLKSKTDQMSMRNW 202

QY 178 GNMWOSNTYLRKQSLSFQVTDSDGRTVYSYDVYPHDMQFGOTFEGGOF 225
 DB 203 GNMWOSNTYLRKQSLSFQVTDSDGRTVYSYDVYPHDMQFGOTFEGGOF 250

RESULT 3
 F84831
 probable expansin [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
 C:Accession: F84831
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanKien, S.E.; Umayam, L.; Tallon, L.;
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; MUID:20083487
 A:Accession: F84831
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-253 <STO>
 A:Cross-references: GB:AEO02093; NID:g2651297; PIDN:AAB87577.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: At2g40610
 A:Map position: 2
 C:Superfamily: expansin

Query Match 77.0%; Score 958.5; DB 2; Length 253;
 Best Local Similarity 74.3%; Pred. No. 1,6e-79;
 Matches 171; Conservative 23; Mismatches 29; Indels 7; Gaps 2;

QY 1 DNGGWHATFFYGADAGTGMGACGYGNLHSGYGLQTAALSTALFNSGQKGCACFEL 60
 DB 26 DYGWGHATFFYGADAGTGMGACGYGNLYSGYGTNTVALSTALFNNGLSCGACFEM 85

QY 61 TCEDDEPWCIPGSIIVRYNLANF-----ALANDNGMCNPLKHPDLAPAFLOIAQYR 115
 DB 86 KGNDDPRKACGSIIV--TATNFCPPNPGLSNDNGMCNPLDHPMAEPALQIAQYR 143

QY 116 GIVPAFRVCEKGGGIRFTINGNPFEDLVLTITNGAGADIRAVSLKSKTDQMSMR 175

DB 144 GIVPSFRVCEKMGKGGIRFTINGNPFEDLVLTITNGAGADIRAVSLKSKTDQMSMR 203

QY 176 NMGWOSNTYLRKQSLSFQVTDSDGRTVYSYDVYPHDMQFGOTFEGGOF 225
 DB 204 NMGWOSNTYLRKQSLSFQVTDSDGRTVYSYDVYPHDMQFGOTFEGGOF 253

RESULT 4
 T09821
 expansin (clone ptxp3) - loblolly pine (fragment)
 C:Species: Pinus taeda (loblolly pine)
 C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 26-May-2000
 C:Accession: T09821
 R:Hutchison, K.W.; Singer, P.B.; Diaz-Sala, C.; Greenwood, M.S.
 submitted to the EMBL Data Library, July 1996
 A:Description: Expansins are conserved in conifers and expressed in response to exoge
 A:Reference number: Z16866
 A:Accession: T09821
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-232 <HUT>
 A:Cross-references: EMBL:U64891; NID:g1778100; PID:g1778101
 A:Experimental source: clone ptxp3
 C:Superfamily: expansin

Query Match 74.5%; Score 927; DB 2; Length 232;
 Best Local Similarity 72.9%; Pred. No. 1e-76;
 Matches 164; Conservative 26; Mismatches 31; Indels 4; Gaps 2;

QY 3 GGMGWHATFFYGADAGTGMGACGYGNLHSGYGLQTAALSTALFNSGQKGCACFEL 62
 DB 8 GGMGWHATFFYGADAGTGMGACGYGNLYSGYGTNTVALSTALFNDGLSCGACFEM 67

QY 63 EDDPWCIPGSIIV---RYNLANFALANDNGMCNPLKHPDLAPAFLOIAQYRAGI 119
 DB 68 NDDPWCIPGSIIVRYNLANFALANDNGMCNPLKHPDLAPAFLOIAQYRAGI 127

QY 120 VAFRRVCEKGGGIRFTINGNPFEDLVLTITNGAGADIRAVSLKSKTDQMSMRNW 179
 DB 128 ILYTRVPCLRKGGIRFTINGNPFEDLVLTITNGAGADIRAVSLKSKTDQMSMRNW 186

QY 180 NMGWOSNTYLRKQSLSFQVTDSDGRTVYSYDVYPHDMQFGOTFEGGOF 224
 DB 187 NMGWOSNTYLRKQSLSFQVTDSDGRTVYSYDVYPHDMQFGOTFEGGOF 251

RESULT 5
 T09826
 expansin (clone ptxp5) - loblolly pine (fragment)
 C:Species: Pinus taeda (loblolly pine)
 C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 26-May-2000
 C:Accession: T09826
 R:Hutchison, K.W.; Singer, P.B.; Diaz-Sala, C.; Greenwood, M.S.
 submitted to the EMBL Data Library, July 1996
 A:Description: Expansins are conserved in conifers and expressed in response to exoge
 A:Reference number: Z16866
 A:Accession: T09826
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-232 <HUT>
 A:Cross-references: EMBL:U64893; NID:g1778104; PID:g1778105
 A:Experimental source: clone ptxp5
 C:Superfamily: expansin

Query Match 74.5%; Score 927; DB 2; Length 232;
 Best Local Similarity 72.9%; Pred. No. 1e-76;
 Matches 164; Conservative 26; Mismatches 31; Indels 4; Gaps 2;

QY 3 GGMGWHATFFYGADAGTGMGACGYGNLHSGYGLQTAALSTALFNSGQKGCACFEL 62
 DB 8 GGMGWHATFFYGADAGTGMGACGYGNLYSGYGTNTVALSTALFNDGLSCGACFEM 67

QY 63 EDDPWCIPGSIIV---RYNLANFALANDNGMCNPLKHPDLAPAFLOIAQYRAGI 119

A:Residues: 1-246 <CHO>
 A:Cross-references: EMBL:U85246, NID:g1815680, PIDN:AA81662.1, PID:g1815681
 A:Experimental source: cv, Pin Gaew 56
 C:Genetics:
 A:Gene: EXP4
 C:Superfamily: expansin

Query Match 71.0%; Score 884; DB 2; Length 246;
 Best Local Similarity 71.2%; Pred. No. 8.7e-73;
 Matches 161; Conservative 23; Mismatches 38; Indels 4; Gaps 2;

QY 3 GGMERGHATFYGGADAGTGMGACGYGNLHSGYGLQTAALSTALFNSGOKGACFELTC 62
 DB 22 GGMOSAHATFYGGADAGTGMGACGYGNLHSGYGLQTAALSTALFNSGOKGACFELTC 81
 QY 63 EDDPEWCIPGSIIV--RYNLANFALANDNGWCNPPKHPDLAPAFLOIAOYRAGIVP 119
 DB 82 DNAGSSCLPGSIVTATNFCPPNYGLPSDDGWCNPPRPHFDMAPAFLOIAOYRAGIVP 141
 QY 120 VAFRRVPCERKGGIRFTINGNPFEDLVLTNNVGAGDIRAVSLKSKTDQWMSRNMGQ 179
 DB 142 VSRFRVPCVKRGGIRFTINGNPFEDLVLTNNVGAGDIRAVSLKSKTDQWMSRNMGQ 200
 QY 180 NMOSNTYLRGOSLSFYVTDSDGRVTVSYDVVPHDMQFGOTFEFGQF 225
 DB 201 NMOSNAFLDGOISLFFVYTAASDGRVTVSNVNAHPGMOFGOTFEFGQF 246

RESULT 10

T03298

Expansin 2 - rice

C:Species: Oryza sativa (rice)

C>Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 21-Jul-2000

C:Accession: T03298

R:Shcherban, T.Y.; Shi, J.; Duracko, D.M.; Gultinan, M.J.; McQueen-Mason, S.J.; Shieh, Proc. Natl. Acad. Sci. U.S.A. 92, 9245-9249, 1995

A:Title: Molecular cloning and sequence analysis of expansins--a highly conserved, multi

A:Reference number: 214894; MUID:96016146

A:Accession: T03298

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-251 <SHC>

A:Cross-references: EMBL:U30477; NID:g1041709; PIDN:AA838074.1; PID:g1041710

C:Genetics:

A:Gene: EXP2

C:Function: induces extension (creep) in plant cell

A:Description: induces extension (creep) in plant cell

C:Superfamily: expansin

Query Match 70.9%; Score 883; DB 2; Length 251;
 Best Local Similarity 70.2%; Pred. No. 1.1e-72;
 Matches 160; Conservative 24; Mismatches 40; Indels 4; Gaps 3;

QY 1 DNGMERGHATFYGGADAGTGMGACGYGNLHSGYGLQTAALSTALFNSGOKGACFELTC 60
 DB 25 DYGSMOSAHATFYGGADAGTGMGACGYGNLHSGYGLQTAALSTALFNSGOKGACFELTC 84
 QY 61 TCEDPEWCIPGSI-IVRYNLAA--NFALANDNGWCNPPKHPDLAPAFLOIAOYRAGI 117
 DB 85 RCDNDGOMCLPGSVTATNLCPPNYGLPSDDGWCNPPRPHFDMAPAFLOIAOYRAGI 144
 QY 118 VPAFRVPCERKGGIRFTINGNPFEDLVLTNNVGAGDIRAVSLKSKTDQWMSRNMGQ 177
 DB 145 VPAFRVPCERKGGIRFTINGNPFEDLVLTNNVGAGDIRAVSLKSKTDQWMSRNMGQ 203
 QY 178 GGNOSNTYLRGOSLSFYVTDSDGRVTVSYDVVPHDMQFGOTFEFGQF 225
 DB 204 GGNOSNTYLRGOSLSFYVTDSDGRVTVSYDVVPHDMQFGOTFEFGQF 251

RESULT 11
 T50654
 expansin EXP1 [imported] - Arabidopsis thaliana (fragment)

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 02-Sep-2000

C:Accession: T50654

R:Shcherban, T.Y.; Shi, J.; Duracko, D.M.; Gultinan, M.J.; McQueen-Mason, S.J.; Shi Proc. Natl. Acad. Sci. U.S.A. 92, 9245-9249, 1995

A:Title: Molecular cloning and sequence analysis of expansins--a highly conserved, mu

A:Reference number: 214894; MUID:96016146

A:Accession: T50654

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-237 <SHC>

A:Cross-references: EMBL:U30476; PIDN:AA838070.1

C:Genetics:

A:Gene: EXP1

C:Function: induces extension (creep) in plant cell walls

A:Description: induces extension (creep) in plant cell walls

C:Superfamily: expansin

C:Keywords: cell wall

Query Match 69.8%; Score 869; DB 2; Length 237;
 Best Local Similarity 69.3%; Pred. No. 1.9e-71;
 Matches 156; Conservative 28; Mismatches 37; Indels 4; Gaps 2;

QY 3 GGMERGHATFYGGADAGTGMGACGYGNLHSGYGLQTAALSTALFNSGOKGACFELTC 62
 DB 12 GGMVNAHATFYGGADAGTGMGACGYGNLHSGYGLQTAALSTALFNSGOKGACFELTC 71
 QY 63 EDDPEWCIPGSIIV--RYNLANFALANDNGWCNPPKHPDLAPAFLOIAOYRAGIVP 119
 DB 72 QNDGKWCCLPGSIVTATNFCPPNYGLPSDDGWCNPPRPHFDMAPAFLOIAOYRAGIVP 131

QY 120 VAFRRVPCERKGGIRFTINGNPFEDLVLTNNVGAGDIRAVSLKSKTDQWMSRNMGQ 179
 DB 132 VAFRRVPCERKGGIRFTINGNPFEDLVLTNNVGAGDIRAVSLKSKTDQWMSRNMGQ 190
 QY 180 NMOSNTYLRGOSLSFYVTDSDGRVTVSYDVVPHDMQFGOTFEFGQF 224
 DB 191 NMOSNTYLRGOSLSFYVTDSDGRVTVSYDVVPHDMQFGOTFEFGQF 235

RESULT 12
 C84444
 probable expansin [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001

C:Accession: C84444
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y. Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487

A:Accession: C84444

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-248 <STO>

A:Cross-references: GB:AE002093; NID:g3461833; PIDN:AA832927.1; GSPDB:GND0139

C:Genetics:

A:Gene: At2g03090

A:Map position: 2

C:Superfamily: expansin

Query Match 69.7%; Score 868; DB 2; Length 248;
 Best Local Similarity 67.8%; Pred. No. 2.5e-71;
 Matches 154; Conservative 31; Mismatches 38; Indels 4; Gaps 2;

QY 2 NGMERGHATFYGGADAGTGMGACGYGNLHSGYGLQTAALSTALFNSGOKGACFELTC 61
 DB 22 DAGVNAHATFYGGADAGTGMGACGYGNLHSGYGLQTAALSTALFNSGOKGACFELTC 81
 QY 62 CEDPEWCIPGSIIV--RYNLANFALANDNGWCNPPKHPDLAPAFLOIAOYRAGIVP 118
 DB 82 CQSDGAWCLFGLATVATNFCPPNYGLPSDDGWCNPPRPHFDMAPAFLOIAOYRAGIVP 141

RESULT 13
D84820

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-257 <STO>
A:Cross-references: GB:AE002093; NID:g2795809; PIDD:AA97125.1; GSPDB:GM00135
C:Gene: At2g39700
A:Map position: 2
C:Superfamily: expansin

Query Match	68.4%;	Score 851;	DB 2;	Length 257;
Best Local Similarity	68.4%;	Pred. No. 9e-70;		
Matches 156;	Conservative 23;	Mismatches 43;		

[illegible]

RESULT 14
02530

Probable expansin F13M22.14 - Arabidopsis thaliana
#Species: Arabidopsis thaliana (mouse-ear cress)
#Date: 05-Mar-1999 #sequence_reviston 05-Mar-1999
#Accession: T02530. #text_change 16-Feb-2001
#Rounsley, S.D.; Kaul, S.; Lilo, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes
submitted to the EMBL data library, June 1998
#Description: Arabidopsis thaliana chromosome II BAC F13M22 genomic sequence.
#Reference number: Z16797
#Accession: T02530

A: Status: translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1..262 <ROU>
A: Cross-references: EMBL:AC004684; NID:g3236234; PID:g3236246
A: Experimental source: Cultivar Colombia
R: Lin. x.: Kaul, S.; Rounsley, S.D.; Shen, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
Muss, D.; Nleman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999.

Query Match	68.3%	Score 850;	DB 2;	Length 2627;
Best Local Similarity	68.9%	Pred. No. 1.1e-69;		
Matches 157; Conservative	23;	Mismatches 42;		Indels

conservative 23; Mismatches 42; Indels 6; Gaps 3;

Db

34 GPMQNAHATFYGGSDASGTMGACGCGYGLNLSQGYGVLTALSTALFNNGFSCGACFEIKC 93

94 TDDPRMCPVGNPSILVTATNFCRPNFAQPSDDGWCNPPREHFDLAMPMLKIGLYRAGI 153

Db

154	VPSYRRVPCRKIGSIRFTVNGFRYENLVLTNVAGACDINGSVSKGSKTD-VWRMSRNW	212
-----	--	-----

Db 213 GQNMQSNVAVLIGQLSFRVTASDRRSSTSMWNPATWQFGQTFSGKNF 260

RESULT 15
T47600

147002
 expansin-like protein - *Arabidopsis thaliana*
 N:Alternate names: protein T22E16.160
 C:Species: *Arabidopsis thaliana* (mouse-ear cress)
 C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 26-May-2000
 C:Accession: T47689
 R:Benes, V.; Wundbach, E.; Drzonek, H.; Ansoerge, W.; Mewes, H.W.; Lemcke, K.; Mayerl
 A:Submitted to the Protein Sequence Database, February 2000
 A:Reference number: 224472

A:Accession: T47689
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-260 <BEN>
A:Cross-references: EMBL.AL132975
A:Experimental source: cultivar Columbia; BAC clone T22E16
C:Genetics:

A;Map position: 3
A;Introns: 52/1; 158/2
A;Note: T22E16.160
C;Superfamily: expansin

Query Match	68.0%	Score 846;	DB 2;	Length 260;
Best Local Similarity	67.5%;	Prod. No. 2.6e-69;		
Matches 154;	Conservative 23;	Mismatches 45;	Indels 6;	Gaps 2

3 GGMWGRHATFFYGGADASTMTGGACGTCGNLHSGCYGLDTAALSTALFNNSGOKGACFELVTC 62

```

63 EDDPEWCIPGSIIVRYNLNF-----ALANDGCGWCPPLKHHEDLAEPFLQIAOYRAGI 11
   ::::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
92 VNDPKWCHPGNPSYFVATATNFCPPNLAAQPSDNGGWCNPPRSHFDLAMPYELKIAEVRAGI 15

```

118 VVAFRRRPCCEKGGCIRFTINGNPFDDVLITNVGGADIRAVSLKSGKTDQWQSMRNW 17
152 VPISYRRACRKSGGIRFTINGHRYFNVLITNVAGADIARTSVKSGSTG-WMSLTRNW 23

178 GQNWQNTYLRGQSLFQYTDSDGRTVSVYDVPHDWFQGFEGGF 225

Sun Oct 13 14:19:05 2002

us-09-896-301-5.rpr

||||| | ||||| | | : : | : ||||| |
Db 211 GQNMOSNAVLVGOISLFRVTSDDRTSTSMNIAPSNMOPGOTFGKNE 258

Search completed: October 11, 2002, 15:01:00
Job time : 11.2915 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 11, 2002, 14:54:05 ; Search time 8.80074 Seconds
(without alignments)
624.466 Million cell updates/sec

Title: US-09-896-301-5

Perfect score: 1245
Sequence: 1 DNGGGERHATFYGGADSG.....SYDVVPHDMQFGQFFGGQF 225

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*

1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*\n2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*\n3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*\n4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*\n5: /cgn2_6/ptodata/2/1aa/PCRTUS.COMB.pep:*\n6: /cgn2_6/ptodata/2/1aa/Backfile1.pep:*\n\nPred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1245	100.0	225	2	US-08-440-517A-5
2	1245	100.0	225	4	US-09-092-160-5
3	962	77.3	227	4	US-09-092-160-7
4	863.5	69.4	226	2	US-08-440-517A-6
5	863.5	69.4	226	4	US-09-092-160-6
6	859	69.0	228	2	US-08-440-517A-2
7	859	69.0	228	4	US-09-092-160-2
8	777.5	62.4	222	2	US-08-440-517A-3
9	777.5	62.4	222	4	US-09-092-160-3
10	734	59.0	225	4	US-09-362-642-2
11	732	58.8	225	2	US-08-845-539-2
12	685.5	55.1	179	2	US-08-845-539-6
13	685.5	55.1	179	4	US-09-362-642-6
14	660.5	53.1	227	2	US-08-440-517A-4
15	660.5	53.1	227	4	US-09-092-160-4
16	609	48.9	167	2	US-08-845-539-4
17	609	48.9	167	4	US-09-362-642-4
18	277	22.2	261	1	US-07-971-096-2
19	277	22.2	261	1	US-08-175-096-2
20	257.5	20.7	246	4	US-08-441-507-21
21	257	20.6	263	1	US-07-971-096-4
22	257	20.6	263	1	US-08-175-096-4
23	257	20.6	263	4	US-08-413-974-6
24	257	20.6	263	4	US-08-434-418-6
25	257	20.6	263	4	US-08-433-288-6
26	257	20.6	263	4	US-08-174-739A-6
27	256.5	20.6	272	4	US-08-441-507-15

28	236.5	19.0	245	4	US-08-441-507-24	Sequence 1, Appl
29	204	16.4	200	4	US-08-441-507-4	Sequence 4, Appl
30	203	16.3	197	4	US-08-441-507-5	Sequence 5, Appl
31	152	12.2	145	4	US-08-413-974-4	Sequence 4, Appl
32	152	12.2	145	4	US-08-434-418-4	Sequence 4, Appl
33	152	12.2	145	4	US-08-433-288-4	Sequence 4, Appl
34	152	12.2	145	4	US-08-174-739A-4	Sequence 4, Appl
35	142	11.4	145	4	US-08-441-507-2	Sequence 2, Appl
36	140	11.2	145	4	US-08-441-507-16	Sequence 16, Appl
37	135	10.8	138	4	US-08-441-507-7	Sequence 7, Appl
38	129.5	10.4	122	4	US-08-441-507-23	Sequence 23, Appl
39	107.5	8.6	54	3	US-08-750-419A-12	Sequence 12, Appl
40	97.5	7.8	86	4	US-08-441-507-8	Sequence 8, Appl
41	85.5	6.9	1147	1	US-08-144-121-3	Sequence 3, Appl
42	85.5	6.9	1147	2	US-08-735-893-3	Sequence 3, Appl
43	85.5	6.9	1165	1	US-08-144-121-2	Sequence 2, Appl
44	85.5	6.9	1165	2	US-08-735-893-2	Sequence 2, Appl
45	83	6.7	463	1	US-08-426-428-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-440-517A-5
; Sequence 5, Application US/08440517A
; Patent No. 5959082
; GENERAL INFORMATION:
; APPLICANT: COSGROVE, DANIEL J.;
; APPLICANT: GUILTINAN, MARK;
; APPLICANT: SCHERBAN, TATYANA;
; APPLICANT: SHI, JUN
; TITLE OF INVENTION: PURIFIED EXPANSIN PROTEINS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INTELLECTUAL PROPERTY OFFICE, THE
; ADDRESS: PENNSYLVANIA STATE UNIVERSITY
; STREET: 113 TECHNOLOGY CENTER
; CITY: UNIVERSITY PARK
; STATE: PENNSYLVANIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 16802-7000
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: NEC 286
; OPERATING SYSTEM: DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/440,517A
; FILING DATE:
; CLASSIFICATION: 530
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 225
; TYPE: AMINO ACID
; TOPOLOGY: UNKNOWN
; US-08-440-517A-5
Query Match 100.0%; Score 1245; DB 2; Length 225;
Best local similarity 100.0%; Pred. No. 3.6e-128;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DNGGGERHATFYGGADSGTGMGACGYNLHSGYGLQTAALSTALFNSGCKGACPEL 60
DB 1 DNGGGERHATFYGGADSGTGMGACGYNLHSGYGLQTAALSTALFNSGCKGACPEL 60
QY 61 TCEDDPENCIGSITIVRYNLANFALANDNGWCNPLKHFDLAPAPLQIAQYRAGIVPV 120
DB 61 TCEDDPENCIGSITIVRYNLANFALANDNGWCNPLKHFDLAPAPLQIAQYRAGIVPV 120
QY 121 AFRVPCPKGGGIRFTINGNPFYDLVLTNNYGAGDIPRAVSLKSKTDQMSNRNNCON 180
DB 121 AFRVPCPKGGGIRFTINGNPFYDLVLTNNYGAGDIPRAVSLKSKTDQMSNRNNCON 180

QY 181 WOSNTYLRGOSLSFOYTDSDGRTVSYDVVPHDMQFGQTEGGOF 225
DB 181 WOSNTYLRGOSLSFOYTDSDGRTVSYDVVPHDMQFGQTEGGOF 225

RESULT 2

US-09-092-160-5
; Sequence 5, Application US/09092160C
; Patent No. 6255466
; GENERAL INFORMATION:
; APPLICANT: Cosgrove, Daniel J
; APPLICANT: McQueen-Mason, Simon
; APPLICANT: Gullitnan, Mark J
; APPLICANT: Shcherban, Tatyana
; APPLICANT: Shi, Jun
; TITLE OF INVENTION: PURIFIED EXPANSIN PROTEINS
; FILE REFERENCE: 1194/1C114US3
; CURRENT APPLICATION NUMBER: US/09/092,160C
; CURRENT FILING DATE: 1998-06-05
; EARLIER APPLICATION NUMBER: 08/440,517
; EARLIER FILING DATE: 1995-05-12
; EARLIER APPLICATION NUMBER: 08/242,090
; EARLIER FILING DATE: 1994-05-12
; EARLIER APPLICATION NUMBER: 08/060,944
; EARLIER FILING DATE: 1993-05-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Arabidopsis
; OTHER INFORMATION: expansin
US-09-092-160-5

Query Match 100.0%; Score 1245; DB 4; Length 225;
Best Local Similarity 100.0%; Pred. No. 3.6e-128;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DNGWGERGHATFFYGADASGTMGACGYGNLHSGYGLQTALSTALFNSGKGCACFEL 60
DB 1 DNGWGERGHATFFYGADASGTMGACGYGNLHSGYGLQTALSTALFNSGKGCACFEL 60
QY 61 TCEDDPEMCIPGSIIVRYNLANFALNDGWCNPLKHFDLAEPALQIAQYRAGI 120
DB 61 TCEDDPEMCIPGSIIVRYNLANFALNDGWCNPLKHFDLAEPALQIAQYRAGI 120
QY 121 AFRVYPCCKGGIRFTINGNPFEDLVLTNNVGAGADIRAVSLKSKTDQMSRMNGN 180
DB 121 AFRVYPCCKGGIRFTINGNPFEDLVLTNNVGAGADIRAVSLKSKTDQMSRMNGN 180
QY 181 WOSNTYLRGOSLSFOYTDSDGRTVSYDVVPHDMQFGQTEGGOF 225
DB 181 WOSNTYLRGOSLSFOYTDSDGRTVSYDVVPHDMQFGQTEGGOF 225

RESULT 3

US-09-092-160-7
; Sequence 7, Application US/09092160C
; Patent No. 6255466
; GENERAL INFORMATION:
; APPLICANT: Cosgrove, Daniel J
; APPLICANT: McQueen-Mason, Simon
; APPLICANT: Gullitnan, Mark J
; APPLICANT: Shcherban, Tatyana
; APPLICANT: Shi, Jun
; TITLE OF INVENTION: PURIFIED EXPANSIN PROTEINS
; FILE REFERENCE: 1194/1C114US3
; CURRENT APPLICATION NUMBER: US/09/092,160C
; CURRENT FILING DATE: 1998-06-05
; EARLIER APPLICATION NUMBER: 08/440,517

EARLIER FILING DATE: 1995-05-12
; EARLIER APPLICATION NUMBER: 08/242,090
; EARLIER FILING DATE: 1994-05-12
; EARLIER APPLICATION NUMBER: 08/060,944
; EARLIER FILING DATE: 1993-05-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: cucumber
; OTHER INFORMATION: expansin
US-09-092-160-7

Query Match 77.3%; Score 962; DB 4; Length 227;
Best Local Similarity 75.0%; Pred. No. 3e-97;
Matches 171; Conservative 26; Mismatches 27; Indels 4; Gaps 2;

QY 1 DNGWGERGHATFFYGADASGTMGACGYGNLHSGYGLQTALSTALFNSGKGCACFEL 60
DB 1 DNGWGERGHATFFYGADASGTMGACGYGNLHSGYGLQTALSTALFNSGKGCACFEL 60
QY 61 TCEDDPEMCIPGSIIV--RYNLANFALNDGWCNPLKHFDLAEPALQIAQYRAGI 117
DB 61 TCEDDPEMCIPGSIIVRYNLANFALNDGWCNPLKHFDLAEPALQIAQYRAGI 120
QY 118 VVYAFRRVPCCKGGIRFTINGNPFEDLVLTNNVGAGADIRAVSLKSKTDQMSRMNGN 177
DB 121 VVYAFRRVPCCKGGIRFTINGNPFEDLVLTNNVGAGADIRAVSLKSKTDQMSRMNGN 179
QY 178 GONWOSNNTYLRGOSLSFOYTDSDGRTVSYDVVPHDMQFGQTEGGOF 225
DB 180 GONWOSNNTYLRGOSLSFOYTDSDGRTVSYDVVPHDMQFGQTEGGOF 227

RESULT 4

US-08-440-517A-6
; Sequence 6, Application US/08440517A
; Patent No. 5959082
; GENERAL INFORMATION:
; APPLICANT: COSGROVE, DANIEL J.;
; APPLICANT: GULLITNAN, MARK;
; APPLICANT: SCHERBAN, TATYANA;
; APPLICANT: SHI, JUN
; TITLE OF INVENTION: PURIFIED EXPANSIN PROTEINS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INTELLECTUAL PROPERTY OFFICE, THE
; ADDRESS: PENNSYLVANIA STATE UNIVERSITY
; STREET: 113 TECHNOLOGY CENTER
; CITY: UNIVERSITY PARK
; STATE: PENNSYLVANIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 16802-7000
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: NEC 286
; OPERATING SYSTEM: DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/440,517A
; FILING DATE:
; CLASSIFICATION: 530
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 226
; TYPE: AMINO ACID
; TOPOLOGY: UNKNOWN
US-08-440-517A-6

Query Match 69.4%; Score 863.5; DB 2; Length 226;

RESULT 6

1 PARENT NO. 9223400
 2 GENERAL INFORMATION:
 3 APPLICANT: Cosgrove, Daniel J
 4 APPLICANT: McQueen-Mason, Simon
 5 APPLICANT: Gullinan, Mark J
 6 APPLICANT: Shcherban, Tatyana
 7 TITLE OF INVENTION: PURIFIED EXPANSIN PROTEINS
 8 FILE REFERENCE: 1194/IC11403
 9 CURRENT APPLICATION NUMBER: US/09/092,160C
 10 EARLIER FILING DATE: 1998-06-05
 11 EARLIER APPLICATION NUMBER: 08/440,517
 12 EARLIER FILING DATE: 1995-05-12
 13 EARLIER APPLICATION NUMBER: 08/242,090
 14 EARLIER FILING DATE: 1994-05-12

EARLIER APPLICATION NUMBER: 08/060,944
EARLIER FILING DATE: 1993-05-12
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 228
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: rice expansin
NAME/KEY: UNSURE
LOCATION: 211
OTHER INFORMATION: Xaa is unknown or other.
US-09-092-160-2

Query Match 69.0%; Score 859; DB 4; Length 228;
Best Local Similarity 69.1%; Pred. No. 5,3e-86;
Matches 154; Conservative 28; Mismatches 37; Indels 4; Gaps 2;

QY 3 GGMERGHATFYGGADASGTMGACGYGNLHSGYGLQTAALSTALFNSGKCGACFELTC 62
DB 3 GGMVNAHATFYGGADASGTMGACGYGNLHSGYGLQTAALSTALFNSGKCGACFELTC 62
QY 63 EDDPEMCIPGSIIVRYNLNLFALNDNGWCNPLKHPDLAPAFLOIAQYRAGIVPAE 119
DB 63 QNDKWCMLPGSIVVATNFCPPNNAIPNAGGCMPPQOHFDSOPFORIAQYRAGIVP 122
QY 120 VAFRRVCEKGGIRFTINGNPFYDLVLTNNVGAGADIRAVSLKSKTDQMSMRNMGQ 179
DB 123 VAFRRVCEKGGIRFTINGNPFYDLVLTNNVGAGADIRAVSLKSKTDQMSMRNMGQ 181
QY 180 NMOSNTYLRGQSLSFQVTDSDGRTVSYDVVPHDMQFGOTFEGG 222
DB 182 NMOSNTYLRGQSLSFQVTDSDGRTVSYDVVPHDMQFGOTFEGG 224

RESULT 8
US-08-440-517A-3
Sequence 3, Application US/08440517A

PATENT NO. 5959082
GENERAL INFORMATION:
APPLICANT: COSGROVE, DANIEL J.;
APPLICANT: GUILTINAN, MARK;
APPLICANT: SHCHERBAN, TATYANA;
TITLE OF INVENTION: PURIFIED EXPANSIN PROTEINS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: INTELLECTUAL PROPERTY OFFICE, THE
STREET: 113 TECHNOLOGY CENTER
CITY: UNIVERSITY PARK
STATE: PENNSYLVANIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 16802-7000
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: NEC 286
OPERATING SYSTEM: DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/440,517A
FILING DATE:
CLASSIFICATION: 530
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 222
TYPE: AMINO ACID
TOPOLOGY: UNKNOWN
US-08-440-517A-3

Query+Match 62.4%; Score 777.5; DB 2; Length 222;

Best Local Similarity 64.1%; Pred. No. 4.1e-77;
Matches 143; Conservative 29; Mismatches 46; Indels 5; Gaps 3;

QY 3 GGMERGHATFYGGADASGTMGACGYGNLHSGYGLQTAALSTALFNSGKCGACFELTC 62
DB 3 GGMVNAHATFYGGADASGTMGACGYGNLHSGYGLQTAALSTALFNSGKCGACFELTC 62
QY 63 EDDPEMCIPGSIIVRYNLNLFALNDNGWCNPLKHPDLAPAFLOIAQYRAGIVPAE 122
DB 63 QNDKWCMLPGSIVVATNFCPPNNAIPNAGGCMPPQOHFDSOPFORIAQYRAGIVP 118
QY 123 RRVPCCKRSGIRFTINGNPFYDLVLTNNVGAGADIRAVSLKSKTDQMSMRNMGQ 182
DB 119 RRVPCCKRSGIRFTINGNPFYDLVLTNNVGAGADIRAVSLKSKTDQMSMRNMGQ 177
QY 183 SNTYLRGQSLSFQVTDSDGRTVSYDVVPHDMQFGOTFEGG 225
DB 178 SNTYLRGQSLSFQVTDSDGRTVSYDVVPHDMQFGOTFEGG 220

RESULT 9
US-09-092-160-3

Sequence 3, Application US/09092160C
PATENT NO. 6235466
GENERAL INFORMATION:
APPLICANT: Cosgrove, Daniel J
APPLICANT: McQueen-Mason, Simon
APPLICANT: Guiltinan, Mark J
APPLICANT: Shcherban, Tatyana
TITLE OF INVENTION: PURIFIED EXPANSIN PROTEINS
FILE REFERENCE: 1194/1C1140S3
CURRENT APPLICATION NUMBER: US/09/092,160C
EARLIER FILING DATE: 1998-06-05
EARLIER APPLICATION NUMBER: 08/440,517
EARLIER FILING DATE: 1998-05-12
EARLIER APPLICATION NUMBER: 08/242,090
EARLIER FILING DATE: 1994-05-12
EARLIER APPLICATION NUMBER: 08/060,944
EARLIER FILING DATE: 1993-05-12
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 222
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: rice expansin
NAME/KEY: UNSURE
LOCATION: (14)..(58)
OTHER INFORMATION: Xaa is unknown or other.
US-09-092-160-3

Query Match 62.4%; Score 777.5; DB 4; Length 222;
Best Local Similarity 64.1%; Pred. No. 4.1e-77;
Matches 143; Conservative 29; Mismatches 46; Indels 5; Gaps 3;

QY 3 GGMERGHATFYGGADASGTMGACGYGNLHSGYGLQTAALSTALFNSGKCGACFELTC 62
DB 3 GGMVNAHATFYGGADASGTMGACGYGNLHSGYGLQTAALSTALFNSGKCGACFELTC 62
QY 63 EDDPEMCIPGSIIVRYNLNLFALNDNGWCNPLKHPDLAPAFLOIAQYRAGIVPAE 122
DB 63 QNDKWCMLPGSIVVATNFCPPNNAIPNAGGCMPPQOHFDSOPFORIAQYRAGIVP 118
QY 123 RRVPCCKRSGIRFTINGNPFYDLVLTNNVGAGADIRAVSLKSKTDQMSMRNMGQ 182
DB 119 RRVPCCKRSGIRFTINGNPFYDLVLTNNVGAGADIRAVSLKSKTDQMSMRNMGQ 177
QY 183 SNTYLRGQSLSFQVTDSDGRTVSYDVVPHDMQFGOTFEGG 225
DB 178 SNTYLRGQSLSFQVTDSDGRTVSYDVVPHDMQFGOTFEGG 220

Query Match	59.0%;	Score 734;	DB 4;	Length 225;
Best Local Similarity	-69.4%;	Pred. No. 2.4e-72;		
Matches 136;	Conservative 20;	Mismatches 32;	Indels 8;	Gaps 4

RESULT 11
US-08-045-539-2
; Sequence 2, Application US/08045539
; Patent No. 5929303

NAME: Bastian, Kevin L.
REGISTRATION NUMBER: 34,774

Query Match	58.8%;	Score 732;	DB 2;	Length 225;
Best Local Similarity	69.4%;	Pred. No. 3.9e-72;		
Matches 136; Conservative	19;	Mismatches 33;	Indels 8;	Gaps 4;

```

QY      1 6 NWGQNNWQSNTYLRQS 191
        11111111 1111
Db      210 NWGQNNWQSNVLTQS 225

```

```

1  :
2  :  APPLICANT:  Bennett, Alan B.
3  :  APPLICANT:  Rose, Jocelyn K.C.
4  :  TITLE OF INVENTION:  Fruit-Specific and Ripening-Regulation
5  :  TITLE OF INVENTION:  Expansin Genes to Control Fruit Texture and Softening
6  :  NUMBER OF SEQUENCES:  8

```

MOLECULE TYPE: protein
US-08-845-539-6

[illegible]

```

RESULT 13
US-09-362-642-6
: Sequence 6, Application US/09362642
: Patent No. 6350935
: GENERAL INFORMATION:
: APPLICANT: Bennett, Alan B.
: APPLICANT: Rose, Jocelyn K.C.
: TITLE OF INVENTION: The Regents of the University of California
: TITLE OF INVENTION: Fruit-Specific and Ripening Regulation Expansin Genes
: FILE REFERENCE: 023070-078210US
: CURRENT APPLICATION NUMBER: US/09/362,642
: CURRENT FILING DATE: 1999-07-27
: NUMBER OF SEQ ID NOS: 8
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 6
: LENGTH: 179
: TYPE: FRT
: ORGANISM: Cucumis melo
US-09-362-642-6

```

Query Match	55.1%;	Score 685.5;	DB 4,	Length 179;
Best Local Similarity	71.0%;	Pred. No. 3.4e-67;		
Matches 130;	Conservative 13;	Mismatches 31;	Indels 9;	Gaps 4

RESULT 14
US-08-440-517A-4
: Sequence 4, Application US/08440517A
: Patent No. 5950082
: GENERAL INFORMATION:
: APPLICANT: COSGROVE, DANIEL J.;
: APPLICANT: GOLLITMAN, MARK;
: APPLICANT: SHOEHBAN, TAYAN;
: APPLICANT: SHI, JUN

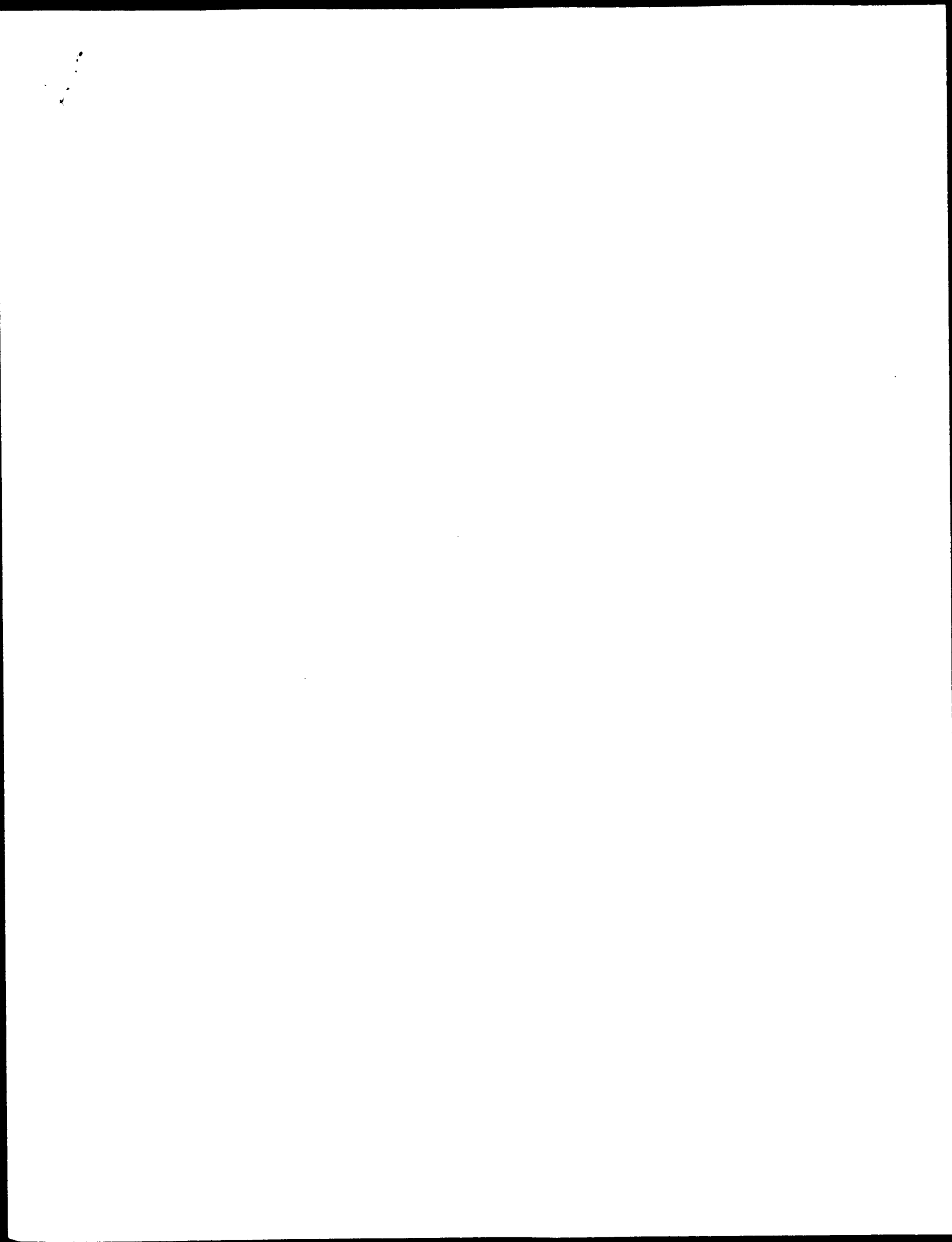
1 TITLE OF INVENTION: PURIFIED EXPANSIN PROTEINS
 2
 3 NUMBER OF SEQUENCES: 6
 4
 5 CORRESPONDENCE ADDRESS:
 6
 7 ADDRESSEE: INTELLECTUAL PROPERTY OFFICE, THE
 8
 9 ADDRESSEE: PENNSYLVANIA STATE UNIVERSITY
 10
 11 STREET: 113 TECHNOLOGY CENTER
 12
 13 CITY: UNIVERSITY PARK
 14
 15 STATE: PENNSYLVANIA
 16
 17 COUNTRY: UNITED STATES OF AMERICA
 18
 19 ZIP: 16802-7000
 20
 21 COMPUTER READABLE FORM:
 22
 23 MEDIUM TYPE: FLOPPY DISK
 24
 25 COMPUTER: NEC 286
 26
 27 OPERATING SYSTEM: DOS
 28
 29 SOFTWARE: WORDPERFECT 5.1
 30
 31 CURRENT APPLICATION DATA:
 32
 33 APPLICATION NUMBER: US/08/440,517A
 34
 35 FILING DATE:
 36
 37 CLASSIFICATION: 530
 38
 39 INFORMATION FOR SEQ. ID NO.: 4:
 40
 41 SEQUENCE CHARACTERISTICS:
 42
 43 LENGTH: 227
 44
 45 TYPE: AMINO ACID
 46
 47 TOPOLOGY: UNKNOWN
 48
 49 US-08-440-517A-4

Query Match	53.1%;	Score 660.5;	DB 2;	Length 227;
Best Local Similarity	57.8%;	Pred. No. 2.5e-64;		
Matches 122;	Conservative 33;	Mismatches 49;	Indels 7;	Gaps 4.

```

RESULT 15
US-09-092-160-4
: Sequence 4, Application US/09092160C
: Patent No. 6255466
:
: GENERAL INFORMATION:
:
: APPLICANT: Cosgrove, Daniel J
: APPLICANT: McQueen-Mason, Simon
: APPLICANT: Gultinan, Mark J
: APPLICANT: Shcherban, Tatyana
: APPLICANT: Shi, Jun
:
: TITLE OF INVENTION: PURIFIED EXPANSIN PROTEINS
:
: FILE REFERENCE: 1194/1C114053
:
: CURRENT APPLICATION NUMBER: US/09/092.160C
:
: CURRENT FILING DATE: 1998-06-05
:
: EARLIER APPLICATION NUMBER: 08/440,517
:
: EARLIER FILING DATE: 1995-05-12
:
: EARLIER APPLICATION NUMBER: 08/242,090
:
: EARLIER FILING DATE: 1994-05-12
:
: EARLIER APPLICATION NUMBER: 08/060,944
:
: EARLIER FILING DATE: 1993-05-12
:
: NUMBER OF SEQ ID NOS: 7
:
: SOFTWARE: PatentIn Ver. 2.1
:
: SEQ ID NO 4
:
: LENGTH: 227
:
: TYPE: PRT
:
: ORGANISM: Artificial Sequence

```

XX Expansins are a novel class of proteins that catalyse the extension
 CC of plant cell walls and the weakening of the hydrogen bonds in pure
 CC cellulose. 2 Expansins (AAR94528 and AAR94529) have been identified in
 CC rice and 3 in Arabidopsis (AAR94530-32). A cDNA clone (AAT13320)
 CC coding for cucumber expansin 29 (AAR94527) has been obtd. Expansins
 CC can be used e.g. in the mfr., de-inking and recycling of paper, in
 CC the textile industry, to aid delignification processes, to alter gel
 CC mechanical strength, etc.

XX Sequence 225 AA;

Query Match 100.0%; Score 1245; DB 17; Length 225;
 Best Local Similarity 100.0%; Pred. No. 4,7e-124;
 Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DNGCMEGHATFFYGADASTGAGCGYGNLHSGYGLQTAALSTALFNSGKGCACFEL 60
 DB 1 DNGCMEGHATFFYGADASTGAGCGYGNLHSGYGLQTAALSTALFNSGKGCACFEL 60
 QY 61 TCEDDEPWCIPGSIIVRYNLANFALANDNGWCNPLKHFDAEPALQIAQYRAGI 120
 DB 61 TCEDDEPWCIPGSIIVRYNLANFALANDNGWCNPLKHFDAEPALQIAQYRAGI 120
 QY 121 AFRVPECEKGGIRFTINGNPFYDLVLTINVGAGDIRAVSLKSGKTDQMSMRNNG 180
 DB 121 AFRVPECEKGGIRFTINGNPFYDLVLTINVGAGDIRAVSLKSGKTDQMSMRNNG 180
 QY 181 WQNTYLRGQSLSFQYTDSDGRIVSYDVYPHDMQFGQTEGGQF 225
 DB 181 WQNTYLRGQSLSFQYTDSDGRIVSYDVYPHDMQFGQTEGGQF 225

RESULT 2

AAR94527 standard; Protein; 227 AA.

XX AAR94527;

DT 08-JUL-1996 (first entry)

DE Cucumber expansin-29.

KW Expansin-29; plant cell wall; cellulose; paper recycling; de-inking;
 KM polysaccharide; cucumber.

OS Cucumis sativus var. Burpee Pickler.

XX A09540262-A.

PN 04-APR-1996.

PF 12-MAY-1994; 94AU-0068320.

PR 12-MAY-1995; 95US-0440517.

PR 12-MAY-1995; 93US-0060944.

XX (PENN-) PENN STATE RES FOUND.

PI Cosgrove DJ, McQueen-Mason S;

DR WPI; 1996-201150/21.

DR N-PSDB; AAT13320.

XX Expansin proteins which alter the mechanical strength of
 PT poly:saccharide(s) - useful in paper mfr. and recycling

XX Claim 7; Page 30; 60pp; English.
 CC Cucumber expansin-29 (AAR94527) is a member of a novel class of
 CC proteins that catalyse the extension of plant cell walls and the
 CC weakening of the hydrogen bonds in pure cellulose. It can be obtd.
 CC by expression of an isolated cDNA clone (see AAT13320) in bacterial or

CC other host cells. Expansin proteins have also been identified in oat
 CC coleoptiles, in Arabidopsis (see AAR94530-32) and in rice (AAR94528-29),
 CC and appear to be broadly distributed throughout the plant kingdom.
 CC Expansins can be used e.g. in the mfr., de-inking and recycling of
 CC paper, in the textile industry, to aid delignification processes, to
 CC alter gel mechanical strength, etc.

XX Sequence 227 AA;

Query Match 77.3%; Score 962; DB 17; Length 227;
 Best Local Similarity 75.0%; Pred. No. 5,9e-94;
 Matches 171; Conservative 26; Mismatches 27; Indels 4; Gaps 2;

QY 1 DNGCMEGHATFFYGADASTGAGCGYGNLHSGYGLQTAALSTALFNSGKGCACFEL 60
 DB 1 DNGCMEGHATFFYGADASTGAGCGYGNLHSGYGLQTAALSTALFNSGKGCACFEL 60
 QY 61 TCEDDEPWCIPGSIIV--RYNLANFALANDNGWCNPLKHFDAEPALQIAQYRAGI 117
 DB 61 TCEDDEPWCIPGSIIVRYNLANFALANDNGWCNPLKHFDAEPALQIAQYRAGI 120
 QY 118 VPIAFRRVPECEKGGIRFTINGNPFYDLVLTINVGAGDIRAVSLKSGKTDQMSMRN 177
 DB 121 VPIAFRRVPECEKGGIRFTINGNPFYDLVLTINVGAGDIRAVSLKSGKTDQMSMRN 179
 QY 178 GQNMQNTYLRGQSLSFQYTDSDGRIVSYDVYPHDMQFGQTEGGQF 225
 DB 180 GQNMQNTYLRGQSLSFQYTDSDGRIVSYDVYPHDMQFGQTEGGQF 227

RESULT 3

AAG25443 standard; Protein; 253 AA.

XX AAG25443;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 29510.

KW Protein identification; signal transduction pathway; metabolic pathway;
 KM hybridisation assay; genetic mapping; gene expression control; promoter;
 termination sequence.

OS Arabidopsis thaliana.

XX EP1033405-A2.

PN 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.

PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147302.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155133.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0156659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 18-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.

PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0133256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137702.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 18-JUN-1999; 99US-0139817.
PR 21-JUN-1999; 99US-0139899.
PR 22-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 23-JUN-1999; 99US-0140695.
PR 24-JUN-1999; 99US-0140823.
PR 28-JUN-1999; 99US-0140825.
PR 29-JUN-1999; 99US-0140891.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 19-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-01452913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149920.
PR 25-AUG-1999; 99US-0150586.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158365.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.

PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 73.38; Score 912; DB 21; Length 241;
Best Local Similarity 71.28; Pred. No. 1.3e-88;
Matches 161; Conservative 30; Mismatches 31; Indels 4; Gaps 2;

OY 3 GGWGRHATPYGADAGTGWGAGCGNLHSGYGLDTALSTLFLMSGKCGACFEFLTC 62
DB 16 GGWLNHATPEYGGDAGTGWGAGCGNLVSGYGTSTALSTALFNNGLSCGSCFEIRC 75
OY 63 EDDPENCIPGSIIV--RYNLANFALNDNGWMCNPLKHFDELAEPFLIOACYRACIYP 119
DB 76 ENDGKWCLPGSIIVTATNFCPPNNALANNNGWMCNPLLEHFDLAQPVFOIRAOYRACIYP 135
OY 120 VAFRRVCEKCGGIRFTINGNPNFDVLTITNVGAGDIRAVSLKSGSTDDQMSKSRMWGQ 179
DB 136 VSYRRVPCRRRGCIIRFTINGHSYENLVLTITNVGAGDVHSAIAIGSRT-VWQMSRMMGQ 194
OY 180 NMOSNTLYLRGOSLSFOYTDSPGRTVSVYDVVPHMORGFQFEFGQF 225
DB 195 NMOSNTLYLRGOSLSFKVYTSQGRTVSVFNAAPGMSYGTFFAGQF 240

RESULT 6
AAC51633
ID AAG51633 standard; protein; 241 AA.

XX AAG51633;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 65551.

KW Protein identification: signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130049.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0142977.
PR 14-JUL-1999; 99US-0143542.
PR 15-JUL-1999; 99US-0143624.
PR 16-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144325.

PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137328.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145931.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0148368.
PR 17-AUG-1999; 99US-0148175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151388.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.

PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 73.3%; Score 912; DB 21; Length 241;
Best Local Similarity 71.2%; Pred. No. 1.3e-88;
Matches 161; Conservative 30; Mismatches 31; Indels 4; Gaps 2;

OY 3 GGMERHATFYGADASTGAGAGYGNLHSGYGLTALSTALFNSGKCGACFELTC 62
DB 16 GGMNNAHTFYGADASTGAGAGYGNLHSGYGLTALSTALFNSGKCGACFELTC 75
OY 63 EDDEMCIPGSIIV--RYNLNANFANDNGMCNPPLEKHEDEPAFLQIAOYRAGIVP 119
DB 76 ENDKRWCLPGSIIVYATNFCEPPNNAANNNGMCNPPLEHEDLQAPYFORIAOYRAGIVP 135
OY 120 VAFRRVPCERGGGIRFTINGNPFYDLVLTINVGAGADIRAVSLKSGKTDQMSRWMGQ 179
DB 136 VSYRRVPCRRRGIRFTINGNPFYDLVLTINVGAGADVHSAIKGSRV--VWQMSRWMGQ 194
OY 180 NMOSNTYRGOSLSPDYVSDGVVSPHDMOFGOPFEGGOF 225
DB 195 NMOSNTYRGOSLSPDYVSDGVVSPHDMOFGOPFEGGOF 240

RESULT 8
AAG06545
ID AAG06545 standard; Protein: 249 AA.

XX AAG06545;
XX
XX 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 3356.

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX
XX

PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 08-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.


```

PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143547.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144337.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144637.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151338.
PR 07-SEP-1999; 99US-0151930.
PR 10-SEP-1999; 99US-0152363.
PR 13-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.

```

```

PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157173.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159684.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Query Match 73.3%; Score 912; DB 21; Length 249;
 Best Local Similarity 71.2%; Pred. No. 1,4e-88;
 Matches 161; Conservative 30; Mismatches 31; Indels 4; Gaps 2;

```

QY 3 GGEHGHATFFYGADSGATMGACGYNLHSGGYGQTALSTALPNSOKGACFELTC 62
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 24 GGINHATFFYGADSGATMGACGYNLHSGGYGQTALSTALPNSOKGACFELTC 62
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 63 EDDPENCIFGSIIV--RYNLAFALANDNGMCNPLKHFPLAEPAFIQIYRAGIYP 119
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 84 ENDGKWCPLGSIYVATNFCPPNNALANNNGMCNPLEHFDLAQVFPRIADYRAGIYP 143
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 120 VAFRRVPCCKGGGIRRTTNGNPFDFVLITNVGGAGDIAVSLKSGSKTDQMSRNNQO 179
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 144 VSTRVPCRRRGIRRTTNGNPFDFVLITNVGGAGDIAVSLKSGSKTDQMSRNNQO 202
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 180 NMQSNLYLRGQSLSFQVTDSDGRITVSYDVPRHMQFGTFEGGQF 225
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 203 NMQSNLYLRGQSLSFQVTDSDGRITVSYDVPRHMQFGTFEGGQF 248
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

RESULT 9
 AAG51632
 ID AAG51632 standard; Protein: 249 AA.
 AC AAG51632;
 XX
 XX
 DT 18-OCT-2000 (First entry)

XX Arabidopsis thaliana protein fragment seq ID NO: 65550.
 DE Protein identification: signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

DE		Arabidopsis thaliana protein fragment SRQ ID NO: 65569.
XX		
KM	Protein identification: signal transduction pathway; metabolic pathway;	
KW	hybridisation assay; genetic mapping; gene expression control; promoter,	
KX	termination sequence.	
OS	Arabidopsis thaliana.	
XX		
PN	EPI033405-A2.	
XX		
PD	06-SEP-2000.	
XX		
Pf	25-FEB-2000; 2000EP-0301439.	
XX		
PR	25-FEB-1999; 99US-0121825.	
PR	05-MAR-1999; 99US-0123180.	
PR	09-MAR-1999; 99US-0123548.	
PR	23-MAR-1999; 99US-0125788.	
PR	25-MAR-1999; 99US-0126264.	
PR	29-MAR-1999; 99US-0126785.	
PR	01-APR-1999; 99US-0127462.	
PR	06-APR-1999; 99US-0128234.	
PR	08-APR-1999; 99US-0128714.	
PR	16-APR-1999; 99US-0129845.	
PR	19-APR-1999; 99US-0130077.	
PR	21-APR-1999; 99US-0130449.	
PR	23-APR-1999; 99US-0130510.	
PR	23-APR-1999; 99US-0130891.	
PR	28-APR-1999; 99US-0131449.	
PR	30-APR-1999; 99US-0132048.	
PR	30-APR-1999; 99US-0132407.	
PR	04-MAY-1999; 99US-0132484.	
PR	05-MAY-1999; 99US-0132485.	
PR	06-MAY-1999; 99US-0132486.	
PR	06-MAY-1999; 99US-0132487.	
PR	07-MAY-1999; 99US-0132863.	
PR	11-MAY-1999; 99US-0134256.	
PR	14-MAY-1999; 99US-0134218.	
PR	14-MAY-1999; 99US-0134219.	
PR	14-MAY-1999; 99US-0134221.	
PR	14-MAY-1999; 99US-0134370.	
PR	18-MAY-1999; 99US-0134768.	
PR	19-MAY-1999; 99US-0134941.	
PR	20-MAY-1999; 99US-0135124.	
PR	21-MAY-1999; 99US-0135353.	
PR	24-MAY-1999; 99US-0135629.	
PR	25-MAY-1999; 99US-0136021.	
PR	27-MAY-1999; 99US-0136382.	
PR	28-MAY-1999; 99US-0136782.	
PR	01-JUN-1999; 99US-0137222.	
PR	03-JUN-1999; 99US-0137528.	
PR	04-JUN-1999; 99US-0137502.	
PR	07-JUN-1999; 99US-0137724.	
PR	08-JUN-1999; 99US-0138094.	
PR	10-JUN-1999; 99US-0138540.	
PR	10-JUN-1999; 99US-0138847.	
PR	14-JUN-1999; 99US-0139119.	
PR	16-JUN-1999; 99US-0139452.	
PR	16-JUN-1999; 99US-0139453.	
PR	17-JUN-1999; 99US-0139492.	
PR	18-JUN-1999; 99US-0139454.	
PR	18-JUN-1999; 99US-0139455.	
PR	18-JUN-1999; 99US-0139456.	
PR	18-JUN-1999; 99US-0139457.	
PR	18-JUN-1999; 99US-0139458.	
PR	18-JUN-1999; 99US-0139460.	
PR	18-JUN-1999; 99US-0139461.	
PR	18-JUN-1999; 99US-0139462.	
PR	18-JUN-1999; 99US-0139463.	
PR	18-JUN-1999; 99US-0139750.	
PR	18-JUN-1999; 99US-0139763.	
PR	21-JUN-1999; 99US-0139817.	

```
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145313.
PR 27-JUL-1999; 99US-0145318.
PR 27-JUL-1999; 99US-0145319.
PR 28-JUL-1999; 99US-0145351.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0156539.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 13-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161993.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 73.3%; Score 912; DB 21; Length 249;
Best Local Similarity 71.2%; Pred. No. 1,4e-88;
Matches 161; Conservative 30; Mismatches 31; Indels 4; Gaps 2;

QY 3 GGERGHATFVGADASGTCGACGYNLHSGCYGLQTAALSTALFNSGOKGACFELTC 62
DB 24 GGTINAHATFVGCGDASGTCGACGYNLHSGCYGLQTAALSTALFNSGOKGACFELTC 83
QY 63 EDDPEWCIPGSIIV--RYNLANFALANDGWCNPEPKHEDLAPFLQIAQYRAGIYP 119
DB 84 ENDGKWCIPGSIIV--RYNLANFALANDGWCNPEPKHEDLAPFLQIAQYRAGIYP 143
QY 120 VAFRRVPCCKGCIPTTNGNPFYFDVLITNNGAGNDIFAVSLKSKTDQMSRNMGO 179
DB 144 VSTRVPCRRRGIRPTINGHSYFNILVTNNGAGDVHSAIKSRT-VWQAKSRNMGO 202
QY 180 NMOSNTYLRGOSLSFOVTDSDGRTVYSYDVPHDMQFGOTFEGGOF 225
DB 203 NMOSNTYLRGOSLSFOVTDSDGRTVYSYDVPHDMQFGOTFEGGOF 248

RESULT 11
AAG06544
ID AAG06544 standard; Protein; 259 AA.
```


PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159844.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161993.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 73.3%; Score 912; DB 21; Length 259;
Best Local Similarity 71.2%; Pred. No. 1.5e-88;
Matches 161; Conservative 30; Mismatches 31; Indels 4; Gaps 2;

QY 3 GGMERHATFTFGADASTGTCAGGTCNLSOGYGLDTALISTVLFNSGKCGACFELTC 62
DB 34 GGMINAHATFTFGADASTGTCAGGTCNLSOGYGLDTALISTVLFNSGKCGACFELTC 93
QY 63 EDDWMCIPGSIIV--RYNLANPALANDNGMGNPLKHFLEPAFLQIAOYRACIIV 119
DB 94 ENDGKWCIPGSIIVTATNFCPPNNAANNNGMGNPLKHFLEPAFLQIAOYRACIIV 153
QY 120 VAFRRVPECKGGIIFFTINGNPPYDVLITNVGAGDIRAVSLGSKTDOMOSMRWQ 179
DB 154 VSYRVRPCRRRGIFFTINGHSYFVLVITNVGAGADVHSAIKGSRF-VWQAMSRWQ 212
QY 180 NMOSNTYLRGOSLSROYTDSGRVVSVDVPMOFGQTEGGO 225
DB 213 NMOSNTYLRGOSLSROYTDSGRVVSVDVPMOFGQTEGGO 258

RESULT 12
ID AAG51631 standard; Protein; 259 AA.
AC AAG51631;
XX 18-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 65549.
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX EP103405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130891.
XX 30-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 06-MAY-1999; 99US-0132487.
XX 07-MAY-1999; 99US-0132863.
XX 11-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 14-MAY-1999; 99US-0134768.
XX 18-MAY-1999; 99US-0134941.
XX 19-MAY-1999; 99US-0135124.
XX 20-MAY-1999; 99US-0135353.
XX 21-MAY-1999; 99US-0135629.
XX 24-MAY-1999; 99US-0136021.
XX 25-MAY-1999; 99US-0136392.
XX 27-MAY-1999; 99US-0136992.
XX 28-MAY-1999; 99US-013782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 16-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.

[illegible]

PR	20-AUG-1999;	99US-01497122
PR	20-AUG-1999;	99US-01497233
PR	20-AUG-1999;	99US-01499029
PR	23-AUG-1999;	99US-0149902
PR	23-AUG-1999;	99US-0149930
PR	25-AUG-1999;	99US-0150556
PR	26-AUG-1999;	99US-0150884
PR	27-AUG-1999;	99US-0151065
PR	27-AUG-1999;	99US-0151066
PR	27-AUG-1999;	99US-0151080
PR	31-AUG-1999;	99US-0151203
PR	31-AUG-1999;	99US-0151438
PR	01-SEP-1999;	99US-0151530
PR	07-SEP-1999;	99US-0152363
PR	10-SEP-1999;	99US-0153070
PR	13-SEP-1999;	99US-0153758
PR	15-SEP-1999;	99US-0154018
PR	16-SEP-1999;	99US-0154039
PR	20-SEP-1999;	99US-0154779
PR	22-SEP-1999;	99US-0155139
PR	23-SEP-1999;	99US-0155486
PR	24-SEP-1999;	99US-0155659
PR	28-SEP-1999;	99US-0156458
PR	29-SEP-1999;	99US-0156596
PR	04-OCT-1999;	99US-0157117
PR	05-OCT-1999;	99US-0157753
PR	06-OCT-1999;	99US-0157865
PR	07-OCT-1999;	99US-0158029
PR	08-OCT-1999;	99US-0158322
PR	12-OCT-1999;	99US-0158363
PR	13-OCT-1999;	99US-0159294
PR	13-OCT-1999;	99US-0159294
PR	13-OCT-1999;	99US-0159295
PR	14-OCT-1999;	99US-0159329
PR	14-OCT-1999;	99US-0159330
PR	14-OCT-1999;	99US-0159331
PR	14-OCT-1999;	99US-0159637
PR	14-OCT-1999;	99US-0159638
PR	14-OCT-1999;	99US-0159584
PR	18-OCT-1999;	99US-0160741
PR	21-OCT-1999;	99US-0160767
PR	21-OCT-1999;	99US-0160768
PR	21-OCT-1999;	99US-0160781
PR	21-OCT-1999;	99US-0160814
PR	21-OCT-1999;	99US-0160815
PR	22-OCT-1999;	99US-0160980
PR	22-OCT-1999;	99US-0160981
PR	22-OCT-1999;	99US-0160989
PR	25-OCT-1999;	99US-0161404
PR	25-OCT-1999;	99US-0161405
PR	25-OCT-1999;	99US-0161359
PR	26-OCT-1999;	99US-0161360
PR	26-OCT-1999;	99US-0161361
PR	26-OCT-1999;	99US-0161920
PR	28-OCT-1999;	99US-0161993
PR	28-OCT-1999;	99US-0161993
PR	29-OCT-1999;	99US-0162142

	Query Match	Similarity	73.3%	Score 91.2	DB 21	Length 259
Best Local	Similarity	71.2%	Pred	No. 1,5e+88		
Matches	161	Conservative	30	Mismatches	31	Indels 4; Gaps 2
QY	3	GSWGRHATFYGGADASGTMGGACGGYGNLHSGYGLQTAALSTALFNSGCKCAGCEHLYIC	62			
Db	34	GMINAHATFYGGDASGTMGGACGGYGNLYSGYGTSTALSTALFNNLSGSGCEHLYIC	93			
QY	63	EEDDEMCIFGSIIV---RYLANFALANDNGMCNPLKHFDAEPFLQIQAYRAGIAP	119			
Db	94	ENDGKKCEFGSIYVATNCPFNALANNNGMCNPELHFDAQVFPQRIQAYRAGIAP	153			
QY	120	FVRRYRCEKGGIIRFTINGNPFEDLVLTITNGGAGDITRAVLSKSGKTQWQSMRNKQO	179			

PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 07-SEP-1999; 99US-0151930.
PR 10-SEP-1999; 99US-0152363.
PR 13-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 23-OCT-1999; 99US-0161404.
PR 23-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 73.3%; Score 912; DB 21; Length 280;
Best Local Similarity 71.2%; Pred. No. 1,6e-88;

Matches 161; Conservative 30; Mismatches 31; Indels 4; Gaps 2;

OY 3 GGNRGHATFYGGADSGATGACGYGNLHSGCYGLQTALSTALFNSQKACAFELTC 62
DB 55 GGNIMHATFYGGCGGACGCGTGNLXSOCTGTALSTALFNNLSCGSCFEIIRC 114
OY 63 EDDPEMCIGESTIV---RYNLANFALANDGCMCNPLKHPDLAPPAFLQIAQYRAGIVP 119

DB 115 ENDGKCLPESIVTATNCPNNMALANNNGMCNPPLEHFDLAPVQRIAQYRAGIVP 174
OY 120 VAFRRVPCCKGKGIKRTFTNCPNEDLYLTNNVGAGDIAVSLKSKTDQMSRNMGQ 179
DB 175 VSTRVPCRRKRGCIKRTFTNHSFNLVLTNNVGAGDVHSAIKSRF-VWQMSRNMGQ 233
OY 180 NMQSNTYLRGQSLSFQVTDSDGRVTVSVDVPHDQFCGTFEGCOF 225
DB 234 NMQSNSYLNQALSFVKVTSDDGRVTVSFMAAPAGNSYQTFAGCOF 279
RESULT 14
AAC23852
ID AAC23852 standard; Protein; 251 AA.
XX AAC23852;
AC
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 27311.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX
PN EP103405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126284.
PR 29-MAR-1999; 99US-0126284.
PR 01-APR-1999; 99US-0126785.
PR 06-APR-1999; 99US-0127462.
PR 08-APR-1999; 99US-0128234.
PR 16-APR-1999; 99US-0128714.
PR 19-APR-1999; 99US-0129845.
PR 21-APR-1999; 99US-0130077.
PR 23-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132047.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135533.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0188847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139452.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 19-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144684.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 23-JUL-1999; 99US-0145276.
 PR 26-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 28-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 02-AUG-1999; 99US-0147038.
 PR 03-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.

PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 18-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152863.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156596.
 PR 29-SEP-1999; 99US-0157117.
 PR 04-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161982.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 69.7%; Score 868; DB 21; Length 251;
 Best Local Similarity 67.8%; Pred. No. 6,7e-84;
 Matches 154; Conservative 31; Mismatches 38; Indels 4; Gaps 2;

QY 2 NCGERCHAFYFGADASTMTGACGCGYLNHSGYGLQTALSTALFNSGOKGACFELT 61
Db 25 DAGVYNNAHAFYGGSDASTMTGACGCGYLNHSGYGLQTALSTALFNSGOKGACFELK 84
QY 62 CEDDPEMCJPGSIIV---RYNLANFALNDNGGMCNPLKHPDLAPALQIQAQRACTV 111
Db 85 CQSDGAMCRLGALIVATNFECPNNALPNNAGGCNPLRHEDLSQVFORIAQRACTV 144
QY 119 PVAFRRVPECKGGGIFRTINGNPPFDLLVTNNGGAGDIANAISLKSGTKDQMSRWG 178
Db 145 PVSIVRRPVMRGGGIFRTINGHSYFNLLVLTNVGGADYHSAVVGSR--RNGQMSRWG 203
QY 179 QNNQSNVTLRGQSLSFQVYLDISGRFVSYVDVPHDQFQGTFFGGQGF 225
Db 204 QNNQSNLLNGQSLFKTKYTAADGRIVVSNIAIPMSVFGQTFGTGRGF 250

RESULT 15	
AA043343	
ID	AA043343 standard; Protein; 251 AA.
XY	

AC AAG43343;

DT 18-OCT-2000 (first entry)
 YY

XX *Aradriopsis* *challana* protein fragment SEQ ID NO: 54163.

KW
KW
KW
hybridisation assay; genetic mapping; gene expression control; promoter termination sequence.

OS Arabidopsis thaliana
xy

PN EP1033405-A2.

PD 06-SEP-2000.

25-FEB-2000; 2000EP-0301439

25-FEB-1999; 99US-0121825.

09-MAR-1999; 99US-0123548.

25-MAR-1999; 99US-0126264.

01-APR-1999; 99US-0127462.

PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128714.

19-APR-1999; 99US-0130077.

23-APR-1999; 99US-0130510.

28-APR-1999; 99US-0131449.

30-APR-1999; 99US-0132407.

05-MAY-1999; 99US-0132485.

06-MAY-1999; 99US-0132487.

11-MAY-1999; 99US-0134256.

14-MAY-1999; 99US-0134219.

14-MAY-1999; 99US-0134370.

19-MAY-1999; 99US-0134941.

21-MAY-1999; 99US-0135353.

27-MAY-1999; 99US-0136021.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

PR	28-MAY-1999	99SUS-0136789
PR	01-JUN-1999	99SUS-0137222
PR	03-JUN-1999	99SUS-0137528
PR	04-JUN-1999	99SUS-0137502
PR	07-JUN-1999	99SUS-0137724
PR	08-JUN-1999	99SUS-0138094
PR	10-JUN-1999	99SUS-0138540
PR	14-JUN-1999	99SUS-0138847
PR	14-JUN-1999	99SUS-0139119
PR	16-JUN-1999	99SUS-0139452
PR	16-JUN-1999	99SUS-0139452
PR	17-JUN-1999	99SUS-0139453
PR	18-JUN-1999	99SUS-0139492
PR	18-JUN-1999	99SUS-0139454
PR	18-JUN-1999	99SUS-0139455
PR	18-JUN-1999	99SUS-0139456
PR	18-JUN-1999	99SUS-0139457
PR	18-JUN-1999	99SUS-0139458
PR	18-JUN-1999	99SUS-0139459
PR	18-JUN-1999	99SUS-0139460
PR	18-JUN-1999	99SUS-0139461
PR	18-JUN-1999	99SUS-0139462
PR	18-JUN-1999	99SUS-0139463
PR	18-JUN-1999	99SUS-0139370
PR	18-JUN-1999	99SUS-01393763
PR	21-JUN-1999	99SUS-0139817
PR	22-JUN-1999	99SUS-0139899
PR	23-JUN-1999	99SUS-0140353
PR	23-JUN-1999	99SUS-0140354
PR	24-JUN-1999	99SUS-0140629
PR	28-JUN-1999	99SUS-0140823
PR	29-JUN-1999	99SUS-0140991
PR	30-JUN-1999	99SUS-0141287
PR	01-JUL-1999	99SUS-0141845
PR	01-JUL-1999	99SUS-0142154
PR	02-JUL-1999	99SUS-0142055
PR	06-JUL-1999	99SUS-0142055
PR	08-JUL-1999	99SUS-0142390
PR	09-JUL-1999	99SUS-0142803
PR	09-JUL-1999	99SUS-0142920
PR	12-JUL-1999	99SUS-0142977
PR	13-JUL-1999	99SUS-0143542
PR	14-JUL-1999	99SUS-0143624
PR	15-JUL-1999	99SUS-0144005
PR	16-JUL-1999	99SUS-0144085
PR	16-JUL-1999	99SUS-0144086
PR	19-JUL-1999	99SUS-0144335
PR	19-JUL-1999	99SUS-0144331
PR	19-JUL-1999	99SUS-0144332
PR	19-JUL-1999	99SUS-0144333
PR	19-JUL-1999	99SUS-0144334
PR	19-JUL-1999	99SUS-0144335
PR	21-JUL-1999	99SUS-0144814
PR	21-JUL-1999	99SUS-0144816
PR	22-JUL-1999	99SUS-0145088
PR	22-JUL-1999	99SUS-0145085
PR	22-JUL-1999	99SUS-0145087
PR	22-JUL-1999	99SUS-0145089
PR	22-JUL-1999	99SUS-0145192
PR	23-JUL-1999	99SUS-0145192
PR	23-JUL-1999	99SUS-0145218
PR	23-JUL-1999	99SUS-0145218
PR	26-JUL-1999	99SUS-0145324
PR	26-JUL-1999	99SUS-0145376
PR	27-JUL-1999	99SUS-0145613
PR	27-JUL-1999	99SUS-0145618
PR	28-JUL-1999	99SUS-0145819
PR	28-JUL-1999	99SUS-0145851
PR	02-AUG-1999	99SUS-0145851
PR	02-AUG-1999	99SUS-0146386
PR	02-AUG-1999	99SUS-0146389
PR	03-AUG-1999	99SUS-0147038
PR	04-AUG-1999	99SUS-0147204

```

PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 18-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 23-OCT-1999; 99US-0161405.
PR 23-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Query Match 69.7%; Score 868; DB 21; Length 251;
Best Local Similarity 67.8%; Pred No. 6.7e-84;
Matches 154; Conservative 31; Mismatches 38; Indels 4; Gaps 2;

QY 2 NGWERSHATFYGADASGTMGACGYNLHSGYGLOTALSTALFNSGCKGACFEIT 61
: || |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 25 DAGVNNHATFYGSDASGTMGACGYNLYSQYGTNTALSTALFNNGLSCGACFEIK 84
: || |||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 62 CEDDPEMCITGSIIV--RYNLNFPALNDNGGMCNPPKHPDLAEPALQIAYRAGIV 118
: || |||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 85 COSDAMCPLPGALIVTATNCPNPNALPNNAGGMCNPPKHPDLISQPVQRIAYRAGIV 144
: || |||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 119 PVARRVPCEKGGIRFTINGNPFYDLVLTITNVGAGDIRAVSLKSKTDQMSRWNG 178
: |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 145 PVSTRVPCARRGSGIRRTINGHSYFNLVLTNVGAGDVHSAVKGSRT-RMQQMSRWNG 203
: |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 179 QNMOSNTYLRCQSLSFOVTDSDGRTVVSXIVYHDQFGQTFEGGOF 225
: |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 204 QNMOSNNILNGLALSFKVTASDGRVVSNNIAPASMSFGQTFEGROF 250
: |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

Search completed: October 11, 2002, 14:56:43
 Job time : 25.9077 secs

QY 60 RCDNDGOMCLPGSVTATATNLCPNNALPNDGGMCPNPPRHFMAEPALQIGYRAGI 119
 DB 85 RCDNDGOMCLPGSVTATATNLCPNNALPNDGGMCPNPPRHFMAEPALQIGYRAGI 144
 QY 120 VPVSYRRVPCVKKGIRFTINGHSYFNLVLTNAGAGDVOSISIGSSGTGMPMSRNWG 179
 DB 145 VPVSYRRVPCVKKGIRFTINGHSYFNLVLTNAGAGDVOSISIGSSGTGMPMSRNWG 204
 QY 180 QNMOSNSYLDGOSLSFOVAASDGRVTYSNNVVPAGMOFGOTFEFGGF 226
 DB 205 QNMOSNSYLDGOSLSFOVAASDGRVTYSNNVVPAGMOFGOTFEFGGF 251

RESULT 2

ID 09461 PRELIMINARY; PRT: 251 AA.
 AC 09461;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DE ALPHA-EXPANSIN.
 GN EXP2.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Eriocaridaceae; Oryzaceae; Oryza.
 RX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21521099; PubMed=11641069;
 RA Lee Y., Choi D., Kende H.;
 RT "Expansins: ever-expanding numbers and functions.";
 RL Curr. Opin. Plant Biol. 4:527-532(2001).
 DR EMBL: AF394544; AAL24480.1;
 SQ SEQUENCE 251 AA; 26674 MW; C45036570FEFE3BD CRC64;

Query Match 97.9%; Score 1230.5; DB 10; Length 251;
 Best local similarity 98.2%; Pred. No. 2.2e-103;
 Matches 223; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 DYSSWOSAHATFFYGGDASGTMGTCGYGNLYSTGY-TNTAALSTVLFNDGACRSCYEL 59
 DB 25 DYSSWOSAHATFFYGGDASGTMGTCGYGNLYSTGYTNTAALSTVLFNDGACRSCYEL 84
 QY 60 RCDNDGOMCLPGSVTATATNLCPNNALPNDGGMCPNPPRHFMAEPALQIGYRAGI 119
 DB 85 RCDNDGOMCLPGSVTATATNLCPNNALPNDGGMCPNPPRHFMAEPALQIGYRAGI 144
 QY 120 VPVSYRRVPCVKKGIRFTINGHSYFNLVLTNAGAGDVOSISIGSSGTGMPMSRNWG 179
 DB 145 VPVSYRRVPCVKKGIRFTINGHSYFNLVLTNAGAGDVOSISIGSSGTGMPMSRNWG 204
 QY 180 QNMOSNSYLDGOSLSFOVAASDGRVTYSNNVVPAGMOFGOTFEFGGF 226
 DB 205 QNMOSNSYLDGOSLSFOVAASDGRVTYSNNVVPAGMOFGOTFEFGGF 251

RESULT 3

ID 094KT7 PRELIMINARY; PRT: 253 AA.
 AC 094KT7;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DE ALPHA-EXPANSIN PRECURSOR.
 GN EXP1.
 OS Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
 OC Panicoidae; Andropogoneae; Zea.
 RX NCBI_TaxID=4577;
 RN [1]
 RP SEQUENCE FROM N.A.

RX MEDLINE=21249159; PubMed=11351085;
 RA Wu Y., Mealey R.B., Cosgrove D.J.;
 RT "Analysis and Expression of the alpha-Expansin and beta-Expansin Gene
 Families in Maize.";
 RL Plant Physiol. 126:222-232(2001).
 DR EMBL: AF332169; AAK56119.1;
 KW SIGNAL.
 FT SIGNAL.
 SQ SEQUENCE 1 20 POTENTIAL.
 253 AA; 26423 MW; 7CDD5390E0104099 CRC64;

Query Match 89.9%; Score 1130.5; DB 10; Length 253;
 Best local similarity 89.4%; Pred. No. 2.4e-94;
 Matches 203; Conservative 7; Mismatches 16; Indels 1; Gaps 1;

QY 1 DYSSWOSAHATFFYGGDASGTMGTCGYGNLYSTGY-TNTAALSTVLFNDGACRSCYEL 59
 DB 27 DYSSWOSAHATFFYGGDASGTMGACGYGNLYSTGYTNTAALSTVLFNDGACRSCYEL 86
 QY 60 RCDNDGOMCLPGSVTATATNLCPNNALPNDGGMCPNPPRHFMAEPALQIGYRAGI 119
 DB 87 RCDNDGOMCLPGSVTATATNLCPNNALPNDGGMCPNPPRHFMAEPALQIGYRAGI 146
 QY 120 VPVSYRRVPCVKKGIRFTINGHSYFNLVLTNAGAGDVOSISIGSSGTGMPMSRNWG 179
 DB 147 VPVSYRRVPCVKKGIRFTINGHSYFNLVLTNAGAGDVOSISIGSSGTGMPMSRNWG 206
 QY 180 QNMOSNSYLDGOSLSFOVAASDGRVTYSNNVVPAGMOFGOTFEFGGF 226
 DB 207 QNMOSNSYLDGOSLSFOVAASDGRVTYSNNVVPAGMOFGOTFEFGGF 253

RESULT 4

ID P93442 PRELIMINARY; PRT: 246 AA.
 AC P93442;
 DT 01-MAY-1997 (TREMBlrel. 03, Created)
 DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE EXPANSIN.
 GN OS-EXP4.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Eriocaridaceae; Oryzaceae; Oryza.
 RX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PIN GAEM 56; TISSUE=INTERNODE;
 RX MEDLINE=97480100; PubMed=9338967;
 RA Cho H.T., Kende H.;
 RT "Expression of expansin genes is correlated with growth in deepwater
 rice.";
 RL Plant Cell 9:1661-1671(1997).
 DR EMBL: U85246; AAB81662.1;
 DR InterPro: IPR000882; Pollen_allergen.
 DR Pfam: PF01357; Pollen_allergen; 1.
 DR PRINTS: PRO1225; EXPANSINFAMLY.
 DR ProDom: PD002179; Pollen_allergen; 1.
 SQ SEQUENCE 246 AA; 25898 MW; CC01823A37BD084F CRC64;

Query Match 87.3%; Score 1097.5; DB 10; Length 246;
 Best local similarity 87.2%; Pred. No. 2.2e-91;
 Matches 197; Conservative 7; Mismatches 21; Indels 1; Gaps 1;

QY 2 YSSWOSAHATFFYGGDASGTMGTCGYGNLYSTGY-TNTAALSTVLFNDGACRSCYEL 60
 DB 21 YSSWOSAHATFFYGGDASGTMGTCGYGNLYSTGYTNTAALSTVLFNDGACRSCYEL 80
 QY 61 CNDNDGOMCLPGSVTATATNLCPNNALPNDGGMCPNPPRHFMAEPALQIGYRAGI 120
 DB 81 CNDNDGOMCLPGSVTATATNLCPNNALPNDGGMCPNPPRHFMAEPALQIGYRAGI 140
 QY 121 VPVSYRRVPCVKKGIRFTINGHSYFNLVLTNAGAGDVOSISIGSSGTGMPMSRNWG 180

DB 141 PVSFRVPCVKKGGIRFTYNGHSYFNLVLTNNVAGAGDVRSISIKSGRTGMPMSRNMGO 200
 OY 181 NMOSNSYLDGQSLSFQVAVSDGRTVTSNNVPAAGMFCGTFEGGOF 226
 DB 201 NMOSNMFLDGQSLSFQVAVSDGRTVTSNNVPAHGMFCGTFEGGOF 246

RESULT 5
 OY46J0 PRELIMINARY: PRT: 246 AA.

AC 0946J0: 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DB ALPHA-EXPANSIN.
 GN EXP4.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21521099; PubMed=11641069;
 RA Lee Y., Choi D., Kende H.;
 RT "Expansins: ever-expanding numbers and functions."
 RL Curr. Opin. Plant Biol. 4:527-532(2001).
 DR EMBL: AF94545; MAF24481.1;
 SQ SEQUENCE 246 AA; 25884 MW; DFAA1224D8A3F9A0 CRC64;

Query Match 87.2%; Score 1096.5; DB 10; Length 246;
 Best Local Similarity 86.7%; Pred. No. 2.7e-91;
 Matches 196; Conservative 8; Mismatches 21; Indels 1; Gaps 1;

OY 2 YSSMSAHATFEYGGDASGTGCTGCGYGNLYSTGY-TNTAALSTVLFPNDGACRSCYELR 60
 DB 21 YGSMQSHATFEYGGDASGTGCTGCGYGNLYSTGY-TNTAALSTVLFPNDGACRSCYELR 80
 OY 61 CDNDGQWCLPGSVYVATNLCPPNYALPNDGWCNPPRPHFMAEPALQIGYVRAGI 120
 DB 81 CDNAGSSCLPGSVYVATNLCPPNYALPNDGWCNPPRPHFMAEPALQIGYVRAGI 140
 OY 121 PVSFRVPCVKKGGIRFTYNGHSYFNLVLTNNVAGAGDVRSISIKSGRTGMPMSRNMGO 180
 DB 141 PVSFRVPCVKKGGIRFTYNGHSYFNLVLTNNVAGAGDVRSISIKSGRTGMPMSRNMGO 200
 OY 181 NMOSNSYLDGQSLSFQVAVSDGRTVTSNNVPAAGMFCGTFEGGOF 226
 DB 201 NMOSNMFLDGQSLSFQVAVSDGRTVTSNNVPAHGMFCGTFEGGOF 246

RESULT 6
 OY9FV30 PRELIMINARY: PRT: 252 AA.
 AC 09FV30: 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DE 01-MAR-2001 (TREMBLrel. 19, Last annotation update)
 DB EXPANSIN.
 GN EXP2.
 OS Festuca pratensis.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
 OC Poaceae; Festuca.
 OX NCBI_TaxID=4608;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=CV. PREPESST; TISSUE=LEAF ELONGATION ZONE;
 RA MEDLINE=21377433; PubMed=11485205;
 RA Reidy B., McQueen-Mason S., Noesberger J., Fleming A.;
 RT "Differential Expression of alpha- and beta-expansin genes in the
 elongating leaf of Festuca pratensis.";

RL Plant Mol. Biol. 46:491-504(2001).
 DR EMBL: AJ276007; CAC06433.1; -
 DR InterPro: IPR000882; Pollen_allergen.
 DR PRINTS: PR01225; EXPANSINFAMILY.
 DR ProDom: PD002179; Pollen_allergen.1.
 SQ SEQUENCE 252 AA; 26544 MW; BE8624E919F42238 CRC64;

Query Match 84.7%; Score 1065; DB 10; Length 252;
 Best Local Similarity 85.1%; Pred. No. 1.9e-88;
 Matches 194; Conservative 10; Mismatches 22; Indels 2; Gaps 2;

OY 1 DYSSMSAHATFEYGGDASGTGCTGCGYGNLYSTGY-TNTAALSTVLFPNDGACRSCYEL 59
 DB 25 DYSSMSAHATFEYGGDASGTGCTGCGYGNLYSTGY-TNTAALSTVLFPNDGACRSCYEL 84
 OY 60 RCDND-GQWCLPGSVYVATNLCPPNYALPNDGWCNPPRPHFMAEPALQIGYVRAGI 118
 DB 85 RCDAAVGTGCRAGSYVATNLCPPNYALPNDGWCNPPRPHFMAEPALQIGYVRAGI 144
 OY 119 IVPVSFRVPCVKKGGIRFTYNGHSYFNLVLTNNVAGAGDVRSISIKSGRTGMPMSRNM 178
 DB 145 IVPVSFRVPCVKKGGIRFTYNGHSYFNLVLTNNVAGAGDVRSISIKSGRTGMPMSRNM 204
 OY 179 GQMSNSYLDGQSLSFQVAVSDGRTVTSNNVPAAGMFCGTFEGGOF 226
 DB 205 GQMSNMFLDGQSLSFQVAVSDGRTVTSNNVPAHGMFCGTFEGGOF 252

RESULT 7
 OY82093 PRELIMINARY: PRT: 254 AA.
 AC 082093: 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DB EXPANSIN.
 GN PA-EXPI.
 OS Prunus armeniaca (Apricot).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eucosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.
 OX NCBI_TaxID=36596;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=BERGERON; TISSUE=MESOCARP PLUS EXOCARP;
 RA Mbeguie A-Mbeguie D., Gomez R.-M., Fils-Lycaon B.;
 RT "Molecular cloning and nucleotide sequence of expansin 1 (Pa-Exp1)
 from apricot fruit.";
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U93167; AAC33529.1; -
 DR InterPro: IPR000882; Pollen_allergen.
 DR Pfam: PF01357; Pollen_allergen.1.
 DR PRINTS: PR01225; EXPANSINFAMILY.
 DR ProDom: PD002179; Pollen_allergen.1.
 SQ SEQUENCE 254 AA; 27264 MW; 88068D75932FD0E1 CRC64;

Query Match 82.1%; Score 1031.5; DB 10; Length 254;
 Best Local Similarity 79.7%; Pred. No. 2.1e-85;
 Matches 181; Conservative 16; Mismatches 23; Indels 1; Gaps 1;

OY 1 DYSSMSAHATFEYGGDASGTGCTGCGYGNLYSTGY-TNTAALSTVLFPNDGACRSCYEL 59
 DB 28 DYSSMSAHATFEYGGDASGTGCTGCGYGNLYSTGY-TNTAALSTVLFPNDGACRSCYEL 87
 OY 60 RCDNDGQWCLPGSVYVATNLCPPNYALPNDGWCNPPRPHFMAEPALQIGYVRAGI 119
 DB 88 RCDNDRPCVKKGGIRFTYNGHSYFNLVLTNNVAGAGDVRSISIKSGRTGMPMSRNM 147
 OY 120 IVPVSFRVPCVKKGGIRFTYNGHSYFNLVLTNNVAGAGDVRSISIKSGRTGMPMSRNM 179
 DB 148 IVPVSFRVPCVKKGGIRFTYNGHSYFNLVLTNNVAGAGDVRSISIKSGRTGMPMSRNM 207
 OY 180 QMSNSYLDGQSLSFQVAVSDGRTVTSNNVPAAGMFCGTFEGGOF 226

Db 208 QNMOSNNYLNQSLSFQVTTSDGRTVTSNNVAPGMOFGOTFSGGOF 254

RESULT 8

09FUM3 PRELIMINARY: PRT: 254 AA.

AC 09FUM3;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE EXPANSIN 1.
GN EXP1 OR EXP2.
OS Prunus avium (cherry), and
OS Prunus cerasus.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.
OX NCBI_TaxID=42229, 140311;

RP SEQUENCE FROM N.A.
RC SPECIES=P.avium;
RA Wu Z., Waters P.A.;
RT "Differential Expression of Expansin Genes Isolated from Sweet Cherry (Prunus avium L.) During Fruit Ripening."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=P.cerasus; TISSUE=RIPENING FRUIT;
RA Yoo S.-D., Gao Z., Cantini C., Loescher W., van Nocker S.;
RT "Coordinated expression of genes encoding expansins and other cell wall-modifying enzymes is associated with pectin-related changes in the cell wall during ripening of cherry (P. cerasus) fruit."
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF297521; AAG13982.1; -
DR EMBL: AF350937; AAK48846.1; -
DR InterPro: IPR000882; Pollen_allergen.
DR Pfam: PF01357; Pollen_allergen.1.
DR PRINTS: PR01225; EXPANSNFAMILY.
DR ProDom: PD002179; Pollen_allergen.1.
SO SEQUENCE 254 AA; 27278 MW; 953A7EB2491FDE01 CRC64;

Query Match 81.8%; Score 1028.5; DB 10; Length 254;
Best Local Similarity 79.3%; Pred. No. 3.9e-85;
Matches 180; Conservative 17; Mismatches 29; Indels 1; Gaps 1;

QY 1 DYSWQSAHATFFYGGDASGTMGTCGYGNLYSTGY-TNTAALSTVLFPNDGACRSCYE 59
DB 28 DYGGMEAHATFFYGGDASGTMGACGYGNLYSGTGNTAALSTLFFNGLSCGSCYE 87
QY 60 RCDNDGQWCLPGSVYVATNLCPPRYALPNDGQWCPNPPHFDMAEPALQIGVYRAGI 119
DB 88 RCNNDPWCPRGSIIVATNFCPPNFAQNDGWCNPLPHFDLAEPAFLQIAQYRAGI 147
QY 120 VPVSYRRVPCVKKGIRFTINGHSYFNLYLVTVNAGPGDVQSYSTGSSGTGMPMRNW 179
DB 148 VPVTFRRVPCVKKGIRFTINGHSYFNLYLVTVNAGGADVHSYISKSGSRGQPMGRNW 207
QY 180 QNMOSNNYLDGQSLSFQVAVSDGRTVTSNNVAPGMOFGOTFSGGOF 226
DB 208 QNMOSNNYLNQSLSFQVTTSDGRTVTSNNVAPGMOFGOTFSGGOF 254

RESULT 9

09SBT1 PRELIMINARY: PRT: 253 AA.

AC 09SBT1;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE EXPANSIN.
GN EXP2.
OS Fragaria ananassa (Strawberry).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Rosales; Rosaceae; Rosoideae; Fragaria.
OX NCBI_TaxID=3747;

RP SEQUENCE FROM N.A.
RC STRAIN=CV, CHANDLER;
RA Civello P.M., Sadehat A., Powell A.L.T., Bennett A.B.;
RT "An expansin gene expressed in ripening strawberry fruit is auxin-independent."
RL Plant Physiol. 12:1273-1279(1999).
DR EMBL: AF159563; AAF21101.1; -
DR InterPro: IPR000882; Pollen_allergen.
DR Pfam: PF01357; Pollen_allergen.1.
DR PRINTS: PR01225; EXPANSNFAMILY.
DR ProDom: PD002179; Pollen_allergen.1.
DR ProSite: PS00626; RCCL2; UNKNOWN_1.
SO SEQUENCE 253 AA; 26887 MW; CE39CF00ADEA1CEF CRC64;

Query Match 81.4%; Score 1023; DB 10; Length 253;
Best Local Similarity 80.3%; Pred. No. 1.2e-84;
Matches 183; Conservative 13; Mismatches 30; Indels 2; Gaps 2;

QY 1 DY-SSWQSAHATFFYGGDASGTMGTCGYGNLYSTGY-TNTAALSTVLFPNDGACRSCYE 58
DB 26 DYGAGWVGHAATFFYGGDASGTMGACGYGNLYSGTGNTAALSTLFFNGLSCGSCYE 85
QY 59 LRCNDGQWCLPGSVYVATNLCPPRYALPNDGQWCPNPPHFDMAEPALQIGVYRAGI 118
DB 86 MRCNDPWCPRGSIIVATNFCPPNFAQNDGWCNPLPHFDLAEPAFLQIAQYRAGI 145
QY 119 VPVSYRRVPCVKKGIRFTINGHSYFNLYLVTVNAGPGDVQSYSTGSSGTGMPMRNW 178
DB 146 VPVTFRRVPCVKKGIRFTINGHSYFNLYLVTVNAGGADVHSYISKSGSRGQPMGRNW 205
QY 179 QNMOSNNYLDGQSLSFQVAVSDGRTVTSNNVAPGMOFGOTFSGGOF 226
DB 206 QNMOSNNYLNQSLSFQVTTSDGRTVTSNNVAPGMOFGOTFSGGOF 253

RESULT 10

P93493 PRELIMINARY: PRT: 232 AA.

AC P93493;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE EXPANSIN (FRAGMENT).
OS Pinus taeda (loblolly pine).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
OX NCBI_TaxID=3352;

RP SEQUENCE FROM N.A.
RC TISSUE=HYPOCOTYL;
RA Hutchison K.W., Singer P.B., Diaz-Sala C., Greenwood M.S.;
RT "Expansins are conserved in conifers and expressed in response to exogenous auxin."
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: U64691; AAB40635.1; -
DR InterPro: IPR000882; Pollen_allergen.
DR Pfam: PF01357; Pollen_allergen.1.
DR PRINTS: PR01225; EXPANSNFAMILY.
DR ProDom: PD002179; Pollen_allergen.1.
FT NON_TER 1

Query Match 81.3%; Score 1021.5; DB 10; Length 232;
Best Local Similarity 79.1%; Pred. No. 1.5e-84;
Matches 178; Conservative 18; Mismatches 28; Indels 1; Gaps 1;

QY 2 YSSWQSAHATFFYGGDASGTMGTCGYGNLYSTGY-TNTAALSTVLFPNDGACRSCYE 60

Db 7 YGWMESAHATFYGGDSAGTGMGACGYGNYLSQGYCTNTALSTALFNDGSCGACTEMR 66
 OY 61 CNDGQWCLPGSVTVATNLCPPNYALPNDGQWGNPPRPHFMAEPALQIGYRAGIV 120
 Db 67 CNDGQWCLPGSVTVATNLCPPNYALPNDGQWGNPPRPHFMAEPALQIGYRAGIV 126
 OY 121 PVSRYRVPCKKGGIRFTINGHSYFNLYVTNNAAGPQVOSVSIKSSGTGQWPMRNMGO 180
 Db 127 PILYTRVPCLRKGGIRFTVNGHSYFNLYVTNNAAGPQVOSVSIKSSGTGQWPMRNMGO 186
 OY 181 NMQNSYLDGQSLSFQYAVSDGRTVTSNNVVPAGMFGQTFEGSQ 225
 Db 187 NMQNSYLDGQSLSFQYAVSDGRTVTSNNVVPAGMFGQTFEGSQ 231

RESULT 11
 O93WP1 PRELIMINARY: PRT: 253 AA.

AC O93WP1. PRELIMINARY: PRT: 253 AA.
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE EXPANSIN.
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
 NCBI_TaxID=3352;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=HYPOCOTYL;
 RX HUTCHISON K.W., SINGER P.B., MCINUS S., DIAZ-SALA C., GREENWOOD M.S.;
 RA "Expansins are conserved in conifers and expressed in hypocotyls in
 RT response to exogenous auxin."
 RL Plant Physiol. 120:827-832(1999).
 DR EMBL: AF085330; AAD47901.1; -
 DR InterPro: IPR000882; Pollen_allergen.1.
 DR Pfam: PF01357; Pollen_allergen.1.
 DR PRINTS: PR01225; EXPANSIN.FAMILY.
 DR ProDom: PD002179; Pollen_allergen.1.
 DR ProDom: PD002179; Pollen_allergen.1.
 SQ SEQUENCE 253 AA; 27150 MW; 2675BBI569BD75D4 CRC64;

Query Match 81.3%; Score 1021.5; DB 10; Length 253;
 Best Local Similarity 79.1%; Pred. No. 1.7e-84;
 Matches 178; Conservative 18; Mismatches 28; Indels 1; Gaps 1;

OY 2 YSSWQSAHATFYGGDSAGTGMGACGYGNYLSQGYCTNTALSTALFNDGSCGACTEMR 60
 Db 28 YGWMESAHATFYGGDSAGTGMGACGYGNYLSQGYCTNTALSTALFNDGSCGACTEMR 87
 OY 61 CNDGQWCLPGSVTVATNLCPPNYALPNDGQWGNPPRPHFMAEPALQIGYRAGIV 120
 Db 88 CNDGQWCLPGSVTVATNLCPPNYALPNDGQWGNPPRPHFMAEPALQIGYRAGIV 147
 OY 121 PVSRYRVPCKKGGIRFTINGHSYFNLYVTNNAAGPQVOSVSIKSSGTGQWPMRNMGO 180
 Db 148 PILYTRVPCLRKGGIRFTVNGHSYFNLYVTNNAAGPQVOSVSIKSSGTGQWPMRNMGO 207
 OY 181 NMQNSYLDGQSLSFQYAVSDGRTVTSNNVVPAGMFGQTFEGSQ 225
 Db 208 NMQNSYLDGQSLSFQYAVSDGRTVTSNNVVPAGMFGQTFEGSQ 252

RESULT 12
 P93492 PRELIMINARY: PRT: 232 AA.
 ID P93492. PRELIMINARY: PRT: 232 AA.
 DT 01-MAY-1997 (Tremblrel. 03, Created)
 DT 01-MAY-1997 (Tremblrel. 03, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE EXPANSIN (FRAGMENT).
 OS Pinus taeda (loblolly pine).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
 NCBI_TaxID=3352;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=HYPOCOTYL;
 RX HUTCHISON K.W., SINGER P.B., DIAZ-SALA C., GREENWOOD M.S.;
 RA "Expansins are conserved in conifers and expressed in response to
 RT exogenous auxin."
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U64890; AAB40634.1; -
 DR InterPro: IPR000882; Pollen_allergen.
 DR Pfam: PF01357; Pollen_allergen.1.
 DR PRINTS: PR01225; EXPANSIN.FAMILY.
 DR ProDom: PD002179; Pollen_allergen.1.
 DR ProDom: PD002179; Pollen_allergen.1.
 FT NON_TER 1
 SQ SEQUENCE 232 AA; 24734 MW; 1340AF483DE0741F CRC64;

Query Match 81.2%; Score 1020.5; DB 10; Length 232;
 Best Local Similarity 78.7%; Pred. No. 1.8e-84;
 Matches 177; Conservative 19; Mismatches 28; Indels 1; Gaps 1;

OY 2 YSSWQSAHATFYGGDSAGTGMGACGYGNYLSQGYCTNTALSTALFNDGSCGACTEMR 60
 Db 7 YGWMESAHATFYGGDSAGTGMGACGYGNYLSQGYCTNTALSTALFNDGSCGACTEMR 66
 OY 61 CNDGQWCLPGSVTVATNLCPPNYALPNDGQWGNPPRPHFMAEPALQIGYRAGIV 120
 Db 67 CNDGQWCLPGSVTVATNLCPPNYALPNDGQWGNPPRPHFMAEPALQIGYRAGIV 126
 OY 121 PVSRYRVPCKKGGIRFTINGHSYFNLYVTNNAAGPQVOSVSIKSSGTGQWPMRNMGO 180
 Db 127 PILYTRVPCLRKGGIRFTVNGHSYFNLYVTNNAAGPQVOSVSIKSSGTGQWPMRNMGO 186
 OY 181 NMQNSYLDGQSLSFQYAVSDGRTVTSNNVVPAGMFGQTFEGSQ 225
 Db 187 NMQNSYLDGQSLSFQYAVSDGRTVTSNNVVPAGMFGQTFEGSQ 231

RESULT 13
 O93XP2 PRELIMINARY: PRT: 253 AA.
 ID O93XP2. PRELIMINARY: PRT: 253 AA.
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE EXPANSIN.
 GN EXPL.
 OS Prunus cerasus.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.
 NCBI_TaxID=140311;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=RIPENING FRUIT;
 RA YOO S.-D., GAO Z., CANTINI C., LOESCHER W., VAN NOCKER S.;
 RT "Coordinated expression of genes encoding expansins and other cell
 RT wall-modifying enzymes is associated with pectin-related changes in
 RT the cell wall during ripening of cherry (P. cerasus) fruit."
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF50936; AAK48845.1; -
 SQ SEQUENCE 253 AA; 26861 MW; 9C24484F14A52CD CRC64;

Query Match 81.1%; Score 1019; DB 10; Length 253;
 Best Local Similarity 79.8%; Pred. No. 2.8e-84;
 Matches 182; Conservative 13; Mismatches 31; Indels 2; Gaps 2;

OY 1 DY-SSWQSAHATFYGGDSAGTGMGACGYGNYLSQGYCTNTALSTALFNDGSCGACTEMR 58
 Db 26 DYGGWEGHATFYGGDSAGTGMGACGYGNYLSQGYCTNTALSTALFNDGSCGACTEMR 85
 OY 59 LRCNDGQWCLPGSVTVATNLCPPNYALPNDGQWGNPPRPHFMAEPALQIGYRAG 118

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 11, 2002, 14:50:30 : Search time 5.83764 Seconds
(without alignments)
1499.000 Million cell updates/sec

Title: US-09-896-301-6
Perfect score: 1257
Sequence: 1 DYSSWQSAHATFYGGDASG.....SNNVYPAGWQPGQTFEGGQF 226

Scoring table: BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 105224 seqs, 38719550 residues
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	232	18.5	246	1	MPCL_CYNDA
2	214.5	17.1	263	1	MPCL_LOLPR
3	207.5	16.5	269	1	MPAL_PHAQ
4	206.5	16.4	265	1	MPH1_HOLLA
5	203	16.1	191	1	MP21_MAIZE
6	201.5	16.0	263	1	MP21_PHLPR
7	175	13.9	263	1	MP21_PRTSA
8	105.5	8.4	418	1	GUN5_TRIRE
9	93	7.4	418	1	GUN2_TRIRE
10	92.5	7.4	646	1	NA95_HUMAN
11	90.5	7.2	333	1	CAT1_MOUSE
12	87	6.9	766	1	DP24_HUMAN
13	86.5	6.9	1246	1	YW2_CAEEL
14	85	6.8	571	1	FLA1_CAMJE
15	84.5	6.7	1176	1	CICB_BACTG
16	84	6.7	3011	1	POLG_HCVI
17	83.5	6.6	513	1	GUX1_TRIVI
18	83.5	6.6	747	1	GUND_CELFI
19	83	6.6	571	1	FLB1_CAMJE
20	82.5	6.6	574	1	FLA3_CAMJE
21	82.5	6.6	574	1	FLB3_CAMJE
22	82.5	6.6	1193	1	LMG2_HUMAN
23	81.5	6.5	1376	1	CRBH_HUMAN
24	81.5	6.5	3695	1	LMA5_HUMAN
25	81	6.4	430	1	SHU2_ECOLI
26	81	6.4	433	1	SHU6_ECOLI
27	81	6.4	442	1	SHU3_ECOLI
28	81	6.4	444	1	SHU4_ECOLI
29	81	6.4	444	1	SHU5_ECOLI
30	81	6.4	456	1	SHU7_ECOLI
31	81	6.4	474	1	SHU1_ECOLI
32	81	6.4	1173	1	TSR1_XENLA
33	80	6.4	677	1	SPB7_DICDI

34	80	6.4	703	1	COAT_SMSV4	p36285 san m1111
35	80	6.4	1429	1	L112_CAEEL	p14585 caenorhabd1
36	80	6.4	3718	1	LMA5_MOUSE	061001 mus musculu
37	79.5	6.3	852	1	CSG_HALNA	p08198 halobacter1
38	79.5	6.3	1385	1	YMS3_CAEEL	p34501 caenorhabd1
39	79.5	6.3	2437	1	NOTC_BRARE	p46530 brachydanio
40	79	6.3	213	1	XYNA_BACSU	p18429 bacillus su
41	78.5	6.2	181	1	GUN_MYTED	p82186 mytilus edu
42	78.5	6.2	513	1	GUX1_TRIRE	p00725 trichoderma
43	78.5	6.2	642	1	NA95_MOUSE	09f017 mus musculu
44	78.5	6.2	1286	1	AIDA_ECOLI	003155 escherichia
45	78	6.2	213	1	XYNA_BACCI	p09850 bacillus ci

ALIGNMENTS

```

RESULT 1
ID MPCL_CYNDA STANDARD: PRT: 246 AA.
AC 004701:
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Major pollen allergen Cyn d 1.
GN CYNDA.
OS Cynodon dactylon (Bermuda grass).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Chloridoideae; Cynodonteae; Cynodon.
OX NCBI_TaxID=28909;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Pollen;
RX MEDLINE=96347957; PubMed=8757211;
RA Smith P.M., Suphloglu C., Griffith I.J., Theriault K., Knox R.B.,
RA Singh M.B.;
RT "Cloning and expression in yeast Pichia pastoris of a biologically
RT active form of Cyn d 1, the major allergen of Bermuda grass pollen ";
RL J. Allergy Clin. Immunol. 98:331-343(1996).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DISEASE: CAUSES GRASS POLLEN ALLERGY.
CC -1- SIMILARITY: BELONGS TO THE LOL P I FAMILY OF ALLERGENS.
CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE EG45 DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE CBD DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: S83343; AAB50734.2; -.
CC HSSP: P43214; IMHO.
CC InterPro: IPR000882; Pollen_allergen.
CC Pfam: PF01357; Pollen_allergen; 1.
CC PRINTS: PR01225; EXPANSINFAMILY.
CC PRODOM: PD002179; Pollen_allergen; 1.
CC PROSITE: PS50843; EXPANSIN_CBD; 1.
CC PROSITE: PS50842; EXPANSIN_EG45; 1.
CC
CC DOMAIN 39 145 EXPANSIN-LIKE EG45.
CC FT DOMAIN 159 240 EXPANSIN-LIKE CBD.
CC FT CARBOHD 9 N-LINKED (GLCNAC...) (POTENTIAL).
CC SO SEQUENCE 246 AA; 26888 MW; 43D8442DBA58322 CRC64;
Query Match 18.5%; Score 232; DB 1; Length 246;
Best Local Similarity 28.3%; Pred. No. 6e-13;
Matches 69; Conservative 37; Mismatches 96; Indels 42; Gaps 10;
OY 3 SWSQSAHATFYGG---GDASGTMGTGCGNLVSTGYTN-TAALSTVLFNDGAACRSCYE 58

```

```

Db 16 SKWLEAFATFGSNRPGAAPDDHGGAGYKDVDPKPPEDGMTACGNEPIFKDGLCCRACYE 75
Qy 59 LRCNDGMCPLPGSVYVTAATNLCPNNALPNDGCMCNPPRRPHDMAEPFLQI----- 113
Db 76 IKCEPEPE-CGGEFVLVKIT-----DKNYEHIAAYHEDLSGKAFGAARKGQE 122
Qy 114 --VYRAGIVPYSRRVPCVKKGIRFTIN---GHSYFNLVLYNVNAGPDVQSVSIR-G 166
Db 123 DKLKAGGLTIQFRVCKYPSGKRTIFHIEKSNHLYLVYXAGDGNIVAVDILKPR 182
Qy 167 SSTGMQPSRNMGNQNMW--SNSYLDGSLSFQAVASDGRVTSNNVYPAGW----- 215
Db 183 DSDDEFIPKSSWGAIWIRIDPKPKPLKG-PFSIRLTSEGAHLVQDDVIPAMWKPDVYTSK 241
Qy 216 -QFG 218
Db 242 LQFG 245

```

RESULT 2

```

ID MPIL_IOLPR STANDARD; PRT; 263 AA.

```

```

AC P1496; P19964;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DE 01-MAR-2002 (Rel. 41, Last annotation update)
DE Pollen allergen Lol p 1 precursor (Lol p 1) (Allergen R7).
OS Lolium perenne (Perennial ryegrass).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Poaceae; Lolium.
OX NCBI_TaxID=4522;
RN 111

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=90375479; PubMed=1697854;
RA Perez M., Ishloka G.Y., Walker L.E., Chesnut R.W.;
RT "cDNA cloning and immunological characterization of the rye grass
RT allergen Lol p 1."
RL J. Biol. Chem. 265:16210-16215(1990).
RN 121

```

```

RP SEQUENCE FROM N.A., AND SEQUENCE OF 24-48.
RC TISSUE=Pollen;
RX MEDLINE=91160716; PubMed=2001733;
RA Griffith I.J., Smith P.M., Pollock J., Theerakulpisut P.,
RA Avoglu A., Davies S., Hough T., Singh M.B., Simpson R.J., Ward L.D.,
RA Knox R.B.;
RT "Cloning and sequencing of Lol p1, the major allergenic protein of
RT rye-grass pollen."
RL FEBS Lett. 279:210-215(1991).
RN 131

```

```

RP SEQUENCE OF 24-53.
RC TISSUE=Pollen;
RX MEDLINE=86242068; PubMed=3718469;
RA Cortam G.P., Moran D.M., Standring R.;
RT "Physicochemical and immunochemical characterization of allergenic
RT proteins from rye-grass (Lolium perenne) pollen prepared by a rapid
RT and efficient purification method."
RL Biochem. J. 234:305-310(1986).
RN 141

```

```

RP SEQUENCE OF 236-263.
RX MEDLINE=89364850; PubMed=2475768;
RA Esch R.E., Klapper D.G.;
RT "Isolation and characterization of a major cross-reactive grass group
RT I allergenic determinant."
RL Mol. Immunol. 26:557-561(1989).
RN 151

```

```

CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DISEASE: CAUSES GRASS POLLEN ALLERGY.
CC -1- SIMILARITY: BELONGS TO THE LOL P 1 FAMILY OF ALLERGENS.
CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE EG45 DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE CBD DOMAIN.

```

```

6C This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

DR EMBL: M57474; AAA63279.1; -
DR EMBL: M57476; AAA63278.1; -
DR PIR: A23341; A23341.
DR PIR: B37881; B37881.
DR PIR: S13614; S13614.
DR HSSP: P43214; 1WHO.
DR InterPro: IPR000882; Pollen.allergen.
DR Pfam: PF01357; Pollen.allergen.1.
DR PRINTS: PR002179; Pollen.allergen.1.
DR PROSITE: PS50843; EXPANSIN_CBD.1.
DR PROSITE: PS50842; EXPANSIN_EG45.1.
KW Allergen; Glycoprotein; Signal; Multigene family.
FT SIGNAL 1 23
FT CHAIN 1 263
FT DOMAIN 61 167
FT DOMAIN 181 262
FT CARBOHYD 32 32
FT VARIANT 68 68
FT VARIANT 177 177
FT VARIANT 210 210
FT VARIANT 246 246
FT CONFLICT 28 28
FT CONFLICT 31 31
FT CONFLICT 48 48
SQ SEQUENCE 263 AA; 28438 MW; 7675896F279C88C9 CRC64;

```

```

Query Match 17.1%; Score 214.5; DB 1; Length 263;
Best Local Similarity 29.1%; Pred. No. 26-11; Mismatches 98; Indels 33; Gaps 11;
Matches 67; Conservative 32;

```

```

Qy 5 WQSAHATFGGAGDASGTM--GGTGGYGNLYSTGYTN-TAALSTVLFNDGACRSCYEELK 61
Db 41 WLDKASTWGTGKPTGAGPKNKGACGCKYKNDKAPFNQMTGCMGPIFKDRGSGSCFEIKC 100
Qy 62 DNDGMCPLPGSVYVTAATNLCPNNALPNDGCMCNPPR-HFDMAEPFLQI-----G 113
Db 101 TKP-ESCSGAEAVVYTT-----DDNE--EPIDAPYHFDLSGHAFGSAKKGEEQN 146
Qy 114 VYRAGIVPYSRRVPCVKKGIRFTINGHSYFN---LVLYNVNAGPGVQSVSIRKGSST 169
Db 147 VRSAGELELQFRVCKYPSGKRTIFHYVERKASPNYALILVYDGDGVAVVDIKRKC 206
Qy 170 G-WQPSRNMGNQNMW--YLDGSLSFQAVASDGRVTSNNVYPAGWQ 216
Db 207 DKWIELKESWGAIWIRIDTPDKLUG-PFTVRYTTEGSKSEFEDVIAPECKK 255

```

RESULT 3

```

ID MPIL_PHAQO STANDARD; PRT; 269 AA.

```

```

AC Q41260;
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 01-MAR-2002 (Rel. 41, Last annotation update)
DE Major pollen allergen Pha a 1 precursor (Pha a 1).
OS Phalaris aquatica (Canary grass).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Poaceae; Phalaris.
OX NCBI_TaxID=28479;
RN 111

```

```

RP SEQUENCE FROM N.A.
RC TISSUE=Pollen;
RX MEDLINE=96105569; PubMed=8564724;
RA Suphlogin C., Singh M.B.;

```


Db 209 DKWIELESGWGVWDTPDKLTG-PPTVRYTEGGTGGAEDYIPBGKADTAYE 263

RESULT 5

MP21_MAIZE STANDARD: PRT: 191 AA.

AC 007154;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Pollen allergen zea m 1 (Zea m 1).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pollen;
RX MEDLINE=94010312; PubMed=8406014;
RA Broadwater A.H., Rubinstein A.L., Chay C.H., Klapper D.G.,
RA Redinger P.A.;
RT "Zea m1, the maize homolog of the allergen-encoding Lol p1 gene of
RT rye grass.";
RL Gene 131:227-230(1993).
CC -1- TISSUE SPECIFICITY: POLLEN TISSUE.
CC -1- DEVELOPMENTAL STAGE: EXPRESSION LOW BEFORE AND HIGH AFTER
CC POLLEN MITOSIS.
CC -1- DISEASE: CAUSES MAIZE POLLEN ALLERGY.
CC -1- SIMILARITY: BELONGS TO THE LOL P 1 FAMILY OF ALLERGENS.
CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE EG45 DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE CBD DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: L14271; AAA33496.1; -
DR PIR: JCI524; JCI524.
DR HSP: P43214; IWHO.
DR MaizeDB: 65840; -
DR InterPro: IPR000882; Pollen_allergen.
DR Pfam: PF01357; Pollen_allergen. 1.
DR PRINTS: PR01225; EXPANSIN_FAM1.
DR ProDom: PD002179; Pollen_allergen; 1.
DR PROSITE: PS50843; EXPANSIN_CBD; 1.
DR PROSITE: PS50842; EXPANSIN_EG45; 1.
KW Allergen; Multigene family.
FT DOMAIN 1 91 EXPANSIN-LIKE EG45.
FT DOMAIN 105 186 EXPANSIN-LIKE CBD.
FT SEQUENCE 191 AA; 21362 MW; 6E2A9DF921C45C63 CRC64;

Query Match 16.1%; Score 203; DB 1; Length 191;
Best Local Similarity 30.6%; Pred. No. 1.4e-10;
Matches 59; Conservative 29; Mismatches 75; Indels 30; Gaps 9;

QY 39 TAAISTVLENDGACRCSCYELRCNDGQWCLPGSVYTAATNLCPNVALPNDGQWGNPP 98
DB 2 TACGNVPIFKKGGCGSCYERCKEPE-CGSPVTFVETIDM--NY-----EPI 47
QY 99 RP-HFDMAEPAF-----LQIGYRAGIVPVSRYRPPVCKKGGIRFTINGHSYFN---- 146
DB 48 APYHFDLSGKAFGLAPGLNDKLRHGGINDYFRYRRCYTPAGQKIVFHEKGCNPNYV 107
QY 147 LVLTIVNAGPDQVSVSIIKGS-STGQPMRNRNQMNQNS--YLDQSLSFQVAASDGR 203
DB 108 AVLVKFAVADGDDIYLMELQKLSAEWKPKMLSMGAIWRMDTAKLKG-PSIRITRSESG 166
QY 204 TVTSSNNVVPAGWQ 216

Db 167 KYIAKDIIPANMR 179

RESULT 6

MP1_PHLPR STANDARD: PRT: 263 AA.

AC P43213;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Pollen allergen Phl p 1 precursor (Phl p 1).
GN PHLPI.
OS Phleum pratense (Common timothy).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
OC Poaceae; Phleum.
OX NCBI_TaxID=15957;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pollen;
RX MEDLINE=95015525; PubMed=7930302;
RA Laffer S., Valente R., Vrtala S., Susani M., van Ree R., Kraft D.,
RA Scheiner O., Duchene M.;
RT "Complementary DNA cloning of the major allergen Phl p 1 from timothy
RT grass (Phleum pratense); recombinant Phl p 1 inhibits IgE binding to
RT group 1 allergens from eight different grass species.";
RL J. Allergy Clin. Immunol. 94:689-698(1994).

CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE LOL P 1 FAMILY OF ALLERGENS.
CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE EG45 DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE CBD DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: X78813; CAA55390.1; -
DR HSP: P43214; IWHO.
DR InterPro: IPR000882; Pollen_allergen.
DR Pfam: PF01357; Pollen_allergen. 1.
DR PRINTS: PR01225; EXPANSIN_FAM1.
DR ProDom: PD002179; Pollen_allergen; 1.
DR PROSITE: PS50843; EXPANSIN_CBD; 1.
DR PROSITE: PS50842; EXPANSIN_EG45; 1.
KW Allergen; Glycoprotein; Signal; Multigene family.
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 263 POLLEN ALLERGEN PHL P 1.
FT DOMAIN 61 167 EXPANSIN-LIKE EG45.
FT DOMAIN 181 262 EXPANSIN-LIKE CBD.
FT CARDHYD 32 32 N-LINKED (GLCNAc... (POTENTIAL).
FT SEQUENCE 263 AA; 28457 MW; 04BA249C17BC048 CRC64;

Query Match 16.0%; Score 201.5; DB 1; Length 263;
Best Local Similarity 26.7%; Pred. No. 2.6e-10;
Matches 63; Conservative 39; Mismatches 101; Indels 33; Gaps 11;

QY 5 WQSAHATYTGGDAGSTM--GGTCGIGNLYSTGYTN-TAALSTVLPNDGACRCSCYELRC 61
DB 41 WLDKASTWGWKPTGAPKPDNGGCGCYKVDKPPFSQMTGCGTPTFKSGRCGSCFETKC 100
QY 62 DNDGQWCLPGSVYTAATNLCPNVALPNDGQWGNPPR-HFDMAEPAFLOIG----- 113
DB 101 TKP-EACSGEPVVVHT-----DDNE--EPYAHFDLSGHAAGAAKAGDQK 146
QY 114 VYRAGIVPVSRYRPPVCKKGGIRFTIN--GHSYFNVLVTNAGPGQVQVSIIKGSST 169
DB 147 LRSAGLELQFRVYKCKYEDGKTYTFHYEKGSGNPNTALLVYVAGDGDVAVVDIKKSG 206

OY 170 G-WQPMRNGMGNOSNS--YLDGOSLSPQVAVSDGRTVTSNNVYPAGWQGFEE 222
DB 207 DKWIELKESMGALWRTDIDKLTG-PFTVRYTTEGKTAEADVIEGKAKDTSTY 261

RESULT 7
ID MPOL_ORYSA STANDARD: PRT: 263 AA.

AC 040638;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 01-MAR-2002 (Rel. 41, Last annotation update)
DR Major pollen allergen Ory s 1 precursor (Ory s I).
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.

RC TISSUE=anther;
RA MEDLINE=96069591; PubMed=7590339;
RA Xu H., Theerakulpisut P., Goulding N., Suphloglu C., Singh M.B.,
Bhalla P.L.;
RT "Cloning, expression and immunological characterization of Ory s 1,
the major allergen of rice pollen.";
RL Gene 164:255-259(1995).

CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN MATURE ANTHERS BUT NOT IN
VEGETATIVE OR OTHER FLORAL TISSUES.
CC -1- DISEASE: CAUSES GRASS POLLEN ALLERGY.
CC -1- SIMILARITY: BELONGS TO THE LOL P I FAMILY OF ALLELGENS.
CC -1- SIMILARITY: CONTAINS 1 EXPANSTIN-LIKE EG45 DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EXPANSTIN-LIKE CBD DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: U31771; AAB6533.1; -
DR HSSP: P43214; 1MHO.
DR InterPro: IPR000882; Pollen_allergen.
DR Pfam: PF01357; Pollen_allergen.1.
DR PRINTS: PR01225; EXPANSINFAMILY.
DR PRODOM: PD002179; Pollen_allergen.1.
DR PROSITE: PS50843; EXPANSTIN_CBD.1.
DR PROSITE: PS50842; EXPANSTIN_EG45.1.
KW Allergen; Glycoprotein; Signal; Multigene family.
FT SIGNAL 1 23
FT CHAIN 1 263
FT MAJOR POLLEN ALLERGEN ORY S 1.
FT DOMAIN 61 164 EXPANSTIN-LIKE EG45.
FT DOMAIN 178 259 EXPANSTIN-LIKE CBD.
FT CARBOHYD 32 32 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SEQUENCE 263 AA; 28497 MW; BICSF24EA398DD60 CRC64;

Query Match 13.9%; Score 175; DB 1; Length 263;
Best Local Similarity 25.7%; Pred. No. 4.7e-08;
Matches 53; Conservative 37; Mismatches 82; Indels 34; Gaps 10;

OY 23 GGTGCGNLYSTGTN-TALSLVLENDGAACSCYELRCNDGCMCLPESVVTATNLC 81
DB 61 GGAGGVDVYKAPFLGNNSGNDPIFDKGGCCSCFEIKCSKP-EACSDKPALIHVTDM- 118
OY 82 PPNVALLNDGCMGNPDP---HFDMAPFAFIQIG---VYRAGIVPVSVRRVPCVKK 133
DB 119 -----NDE-----PLAAVHFDLISGLAMKDGDEELRAGAGIIDJFRHYKCKYPA 163
OY 134 GIREFTIN---GHISYFNLVLTNAGPCDVQSVISK-CSSGTGQPMRNGMGNOSNS-- 186

DB 164 DTKITFHEKASNPVYALTLVKKVYAGDGVVEIKEKSGSEEMKALKESMGALWRTDIDKLTG-PFTVRYTTEGKTAEADVIEGKAKDTSTY 223
OY 187 YLDGOSLSPQVAVSDGRTVTSNNVYP 212
DB 224 PLKG-PFSVRYTTEGARSADALP 248

RESULT 8
ID GUN5_TRIE STANDARD: PRT: 242 AA.

AC P4317;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE 15-DEC-1998 (Rel. 37, Last annotation update)
DR Endoglucanase V precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase V)
DE (Cellulase V) (EG V).
GN Egl5.
OS Trichoderma reesei (Hypocrea jecorina).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreales; Hypocreaceae; Hypocrea.

OX NCBI_TaxID=51453;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=QM9414 / RUT C-30;
RX MEDLINE=95073308; PubMed=7984103;
RA Saloheimo A., Hennisat B., Hoffren A.-M., Telemann O., Penttilae M.;
RT "A novel, small endoglucanase gene, egl5, from Trichoderma reesei
isolated by expression in yeast.";
RL Mol. Microbiol. 13:219-228(1994).

CC -1- CATALYTIC ACTIVITY: Endohydrolase of 1,4-beta-D-glucosidic
linkages in cellulose.
CC -1- SIMILARITY: BELONGS TO CELLULOSE FAMILY K (FAMILY 45 OF GLYCOSYL
HYDROLASES).
CC -1- SIMILARITY: CONTAINS 1 FUNGAL-TYPE CELLULOSE-BINDING DOMAIN (CHU).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: Z3381; CAA83846.1; -
DR HSSP: P00725; 2CBH.
DR InterPro: IPR000254; CBD_fungal.
DR Pfam: PF00734; CBD_1.1.
DR Pfam: PF02015; Glyco_hydro_45.1.
DR PRODOM: PD001821; CBD_fungal.1.
DR SMART: SM00236; FCBP.1.
DR PROSITE: PS00562; CBD_FUNGAL.1.
DR PROSITE: PS50842; EXPANSTIN_EG45.1.
DR PROSITE: PS01140; GLYCOSYL_HYDROL_F45.1.
DR PROSITE: PS01140; GLYCOSYL_HYDROL_F45.1.
KW Cellulose degradation; Hydrolase; Glycosidase; Signal.
FT SIGNAL 1 17
FT CHAIN 18 242
FT ACT_SITE 27 27
FT ACT_SITE 134 134
FT ACT_SITE 182 182
FT DOMAIN 183 205
FT DOMAIN 206 242
FT CARBOHYD 182 182
FT DISULFID 213 230
FT DISULFID 224 240
FT SEQUENCE 242 AA; 24411 MW; CC033FC51326C71D CRC64;

Query Match 8.4%; Score 105.5; DB 1; Length 242;
Best Local Similarity 26.9%; Pred. No. 0.037;
Matches 43; Conservative 18; Mismatches 74; Indels 53; Gaps 11;

OY 3 SSWQSHAFYGG-----GDASGTMGTCGYGN-LYSTGTNTALSLVLEND----- 49

Db 16 SAYATTTTRYDGGEGACGCGSSGAPFMOLGNGVY-----TAAGSQAFLPTAGASW 69
 QY 50 -GAACRSCYELR-----CDNDGQMCPLPG-SVTATATNLCPPNALPNDGCGNPPRP- 100
 Db 70 CGAGCGKCYQLTSTGQAFCCSSCGTGAAGOSIYVNTNLC-----NNGNQMC-----PV 120
 QY 101 -----HFDMAEPALQIGYRAGIVPVSRYRVP 129
 Db 121 VGTGNYGYSYHFDI-----MAONEIFGDNVY-VDPEFLIAC 155

RESULT 9
 GUN2-TRIPE STANDARD; PRT; 418 AA.
 ID GUN2-TRIPE STANDARD; PRT; 418 AA.
 AC P07982;
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Endoglucanase EG-II precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase) (cellulase).
 GN EGL2 OR EGLI.
 OS Trichoderma reesei (Hypocrea jecorina).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Hypocreales; Hypocreaceae; Hypocrea.
 OX NCBI_TaxID=51453;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-VTT-D-80133;
 RX MEDLINE=88255850; PubMed=3384334;
 RA Saloheimo M., Lehtovaara P., Penttilae M., Teeri T.T., Staahlberg J.,
 Johansson G., Pettersson G., Claysens M., Tomme P., Knowles J.K.C.;
 RT "EGII, a new endoglucanase from Trichoderma reesei: the
 RT characterization of both gene and enzyme.";
 RL Gene 63:11-21(1988).
 RN [2]
 RP ACTIVE SITE GLU-350.
 RX MEDLINE=93131031; PubMed=8093602;
 RA Macarron R., van Beeumen J., Henriessat B., de la Mata I.,
 RA Claeysens M.;
 RT "Identification of an essential glutamate residue in the active site
 RT of endoglucanase III from Trichoderma reesei.";
 RL FEBS Lett. 316:137-140(1993).
 CC
 CC -1- FUNCTION: THE BIOLOGICAL CONVERSION OF CELLULOSE TO GLUCOSE
 CC GENERALLY REQUIRES THREE TYPES OF HYDROLYTIC ENZYMES:
 CC (1) EXOCELLULOBIOLYASES WHICH CUT INTERNAL BETA-1,4-GLUCOSIDIC BONDS;
 CC (2) EXOCELLULOBIOLYASES THAT CUT THE DISSACCHARIDE CELLULOSE
 CC FROM THE NONREDUCING END OF THE CELLULOSE POLYMER CHAIN;
 CC (3) BETA-1,4-GLUCOSIDASES WHICH HYDROLYZE THE CELLULOSE AND OTHER
 CC SHORT CELLO-OLIGOSACCHARIDES TO GLUCOSE.
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
 CC linkages in cellulose.
 CC -1- SIMILARITY: BELONGS TO CELLULASE FAMILY A (FAMILY 5 OF GLYCOSYL
 CC HYDROLASES).
 CC -1- SIMILARITY: CONTAINS 1 FUNGAL-TYPE CELLULOSE-BINDING DOMAIN (CBD).
 CC -1- CAUTION: WAS ORIGINALLY CALLED ENDOGLUCANASE EG-II.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL collaboration
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: M19373; AAA34213.1; -
 CC PIR: S28372; S28372.
 CC HSSP: P00725; ZCBH.
 CC InterPro: IPR000254; CBD_fungal.
 CC InterPro: IPR001547; glyco_hydro_f5.
 CC Pfam: PF00734; CBD_1; 1.
 CC Pfam: PF00150; cellulase_1.
 CC ProDom: PD001821; CBD_fungal; 1.
 CC SMART: SM00236; tCBD; 1.

DR PROSITE: PS00562; CBD_FUNGAL; 1.
 DR PROSITE: PS00559; GLYCOSYL_HYDROL_F5; 1.
 KW Cellulose degradation; Hydrolase; Glycosidase; Glycoprotein; Signal.
 FT SIGNAL 1 21
 FT CHAIN 22 418
 FT DOMAIN 22 57
 FT DOMAIN 58 91
 FT DOMAIN 92 418
 FT MOD RES 22 22
 FT CARBOHYD 124 124
 FT DISULFID 29 46
 FT DISULFID 40 56
 FT ACT_SITE 239 239
 FT ACT_SITE 350 350
 SQ SEQUENCE 418 AA; 44227 MW; 264A92D552374A9B CRC64;

Query Match 7.4%; Score 93; DB 1; Length 418;
 Best Local Similarity 23.3%; Pred. No. 0.76;
 Matches 50; Conservative 21; Mismatches 72; Indels 72; Gaps 11;

QY 8 AAATFYGGDASGTMGTCGYGMLYSTGYTNALSTVFNDAACRSCYELRCNDGOW 67
 Db 11 AASIIYGAVAADQTVWGQC--GGIGWSGPTNCA-----PGSACSTLNPYYAQ----- 55
 QY 68 CLPGSVTATNLCPPNVALPNDGCGCNPDPHFMAEPALQIGYRAGIVPVSRYR 127
 Db 56 CIGCATTTT-TSTRPSPGPTTTTTRATSTSSSP-----PTS----- 90
 QY 128 PCVKKKGIRFTNGHSFYMLVLTNVAG-----PGDVQSVSIRKSGSTGOMPSSRMGOW 182
 Db 91 -----SGVFEFA-----GVNIAGDFGCTGTGTCV---TSKVPPLKNTFTG--- 127
 QY 183 OSNSTYLDGQSLSPQAVSDGRVTSNNVYPAGMOF 217
 Db 128 -SNNPDPGIGQMCHFYVNEGDMTIFR---LPVGMQY 158

RESULT 10
 NA95_HUMAN
 ID NA95_HUMAN STANDARD; PRT; 646 AA.
 AC Q9UAX6; Q9UAX6; Q9NR00; Q94792;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Neighbor of A-kinase anchoring protein 95 (Homologous to AKAP95
 DE protein) (HA95) (Helicase A-binding protein 95) (HA95).
 GN NAKAP95.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fetal brain;
 RX MEDLINE=20163068; PubMed=10697960;
 RA Seki N., Ueki N., Yano K., Saito T., Masuho Y., Muramatsu M.-A.;
 RT "cDNA cloning of a novel human gene NAKAP95, neighbor of A-kinase
 RT anchoring protein 95 (AKAP95) on chromosome 19p13.11-p13.12 region.";
 RL J. Hum. Genet. 45:31-37(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20222332; PubMed=10761695;
 RA Orstavik S., Elde T., Collas P., Han I.O., Taaken K., Kieff E.,
 RA Jahnson T., Skalhogg B.S.;
 RT "Identification, cloning and characterization of a novel nuclear
 RT protein, HA95, homologous to A-kinase anchoring protein 95.";
 RL Biol. Cell 92:27-37(2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=T-cell lymphoma, and placenta;
 RX MEDLINE=20347256; PubMed=10748171;
 RA Westberg C., Yang J.-P., Tang H., Reddy T.R., Wong-Skaal F.;
 RT "A novel shuttle protein binds to RNA helicase A and activates the

retroviral constitutive transport element";
 J. Biol. Chem. 275:21396-21401(2000).
 [4]
 RP SEQUENCE OF 1-358 FROM N.A.
 RC TISSUE=Fetal brain:
 RX MEDLINE=99068504; PubMed=9853615;
 RA Ueki N., Oda T., Kondo M., Yano K., Noguchi T., Muramatsu M.-A.;
 RT "Selection system for genes encoding nuclear-targeted proteins";
 RL Nat. Biotechnol. 16:1338-1342(1998).
 CC -1- FUNCTION: COULD PLAY A ROLE IN CONSTITUTIVE TRANSPORT ELEMENT
 (CTE)-MEDIATED GENE EXPRESSION. DOES NOT SEEM TO BE IMPLICATED IN
 THE BINDING OF REGULATORY SUBUNIT II OF PKA.
 CC -1- SUBUNIT: BINDS TO THE C-TERMINAL OF RNA HELICASE A.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR AT STEADY STATE BUT SHUTTLES BETWEEN
 THE NUCLEUS AND CYTOPLASM.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AB025905; BAA85003.1; -;
 DR EMBL; AJ243467; CAB65092.1; -;
 DR EMBL; AF199414; NAF86048.1; -;
 DR EMBL; AB015332; BAA34791.1; ALT_INT.
 DR InterPro: IPR000822; Znf_C2H2; 1.
 DR SMART; SM00355; Znf_C2H2; 1.
 KW Nuclear protein; Zinc-finger.
 FT DOMAIN 274 279
 FT 362 364
 FT 391 413
 FT 484 507
 FT 41 52
 FT 602 608
 FT 589 597
 FT 100 100
 FT 189 189
 FT 351 358
 FT 458 458
 FT 554 596
 FT 610 640
 FT 646 AA: 71648 MW: EEB855A81BB06585 CRC64;
 SQ SEQUENCE
 Query Match 7.4%; Score 92.5; DB 1; Length 646;
 Best Local Similarity 23.6%; Pred. No. 1.3;
 Matches 60; Conservative 29; Mismatches 108; Indels 57; Gaps 13;
 2 YSSWQSHATFYGGCD-ASGTMGCTCCYGNLYSTGYTNTAALSTVLFPDCAACRSCYEL- 59
 18 YSDTSAQPTCDYGTWNSGTNGYGYGYGCDNTTN-----YGYMATSHSMEP 72
 60 RCDNDQWCLPBGSVYATATLCPNPYALPNDGCMCPRRPHDMAEPFLQGYVRAGI 119
 73 SSDTNANTASASGASADSV-LSRINORL-----DMVPHL---ETDMQGGVYGG- 118
 120 VPSYRVNCPVKKGIFETI-----NGHSYFNVLVTNACPGDVQS---VSTKGS 168
 119 -----GERYDYDESCDSRAVYSEKDYLSRGYDYSELDEPMAMAEQYDAYRDFRRKAND 174
 169 T-----GW-----QPMRNMGQNMQNSYLDGSLSFQVAVSDGRVTYSNNVP- 212
 175 TFGPRAGCMARDARSGRPMASGVGKRWMEDEPMGARGCMS---GASRLPSLFSQNIIPERG 231
 213 --AGMFGGTFEGG 224
 232 MFGMGCGCAFGCG 245

RESULT 11
 CAT_MOUSE
 ID CAT_MOUSE STANDARD; PRT: 333 AA.
 AC 09R014; 09W51;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cathepsin J, a novel murine cysteine protease of the papain family
 GN CTSJ OR CTSJ
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Embryo;
 RX MEDLINE=99456833; PubMed=10526153;
 RA Tislar K., Deussing J., Peters C.;
 RT "Cathepsin J, a novel murine cysteine protease of the papain family
 with a placenta-restricted expression";
 RL FEBS Lett. 459:299-304(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Placenta;
 RA Sol-Church K., French J., Troeber D., Mason R.W.;
 RT "Cloning of a mouse cysteine protease";
 RL Submitted (JUN-1999) to the EMBL/Genbank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Lysosomal (Potential).
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN PLACENTA.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1; ALSO KNOWN AS THE
 PAPAINE FAMILY OF THIOL PROTEASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF136272; AAF13142.1; -;
 DR EMBL; AF158182; AAD41898.1; -;
 DR HSSP; P07711; 1CJL.
 DR MEROPS; C01.038; -;
 DR MGD; MGI:1349426; Ctsj.
 DR InterPro: IPR000668; Peptidase_C1.
 DR InterPro: IPR000169; Thiolprote_act_site.
 DR Pfam; PF00112; Peptidase_C1; 1.
 DR PRINTS; PR00705; PAPAINE.
 DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
 DR PROSITE; PS00639; THIOL_PROTEASE_HIS; 1.
 DR PROSITE; PS00640; THIOL_PROTEASE_ASN; 1.
 KW Hydrolase; Thiol protease; Glycoprotein; Lysosome; Zymogen; Signal.
 FT SIGNAL 1 17
 FT PROPEP 18 112
 FT CHAIN 113 333
 FT ACT_SITE 137 137
 FT ACT_SITE 275 275
 FT ACT_SITE 299 299
 FT DISULFID 134 177
 FT DISULFID 168 210
 FT DISULFID 268 321
 FT CARBOHYD 71 71
 FT CARBOHYD 216 216
 FT CARBOHYD 220 220
 FT CARBOHYD 267 267
 FT CONFLICT 42 42
 FT CONFLICT 42 42
 SQ SEQUENCE 333 AA: 37147 MW: F9A8FP1DSAI3B721 CRC64;
 Query Match 7.2%; Score 90.5; DB 1; Length 333;
 Best Local Similarity 25.0%; Pred. No. 0.98;
 Matches 57; Conservative 16; Mismatches 84; Indels 71; Gaps 14;

```

OY 1 DYSSW-QSAHAT-----FYGGDASGMMGTGCGNLVSTGNTAALSTVL 46
DB 115 DYKDMREGGIVTPVRNCKGCCSCNAFAAAGAIIEGOMENKT--GNTITPLSVONLIDCKSTV 172
OY 47 FNDGAACRS-----CYELRCNDNDGOWCLPGSVTVTATNLCPPNYALPNDDGMCNPPRRP 101
DB 173 GNKG--COSGTAHQAFEVVLNKG-----LEAEAT-----YPEGKDG---PCRYR 213
OY 102 FDMAE-----PAFLQIGYVRAGIVPV-----SYRAVPCVKKGIGFTINGH 142
DB 214 SENASANTIDVNLPPNMLYLVAVASIGVPSAIDAIDASHDSFR---FYNGGCIYEPNCS 269
OY 143 SYF--NLVLTNNVAGPGDPVQSVISGSGTGQPMRSRMGMQMSNSYL 188
DB 270 SYFVNHAVLVYVGSEGDV-----KDGNNYVL-INKSMGEEMGNMGV 311

RESULT 12
DEPP4_HUMAN
ID DEPP4_HUMAN STANDARD; PRT; 766 AA.
AC P27487;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Dipeptidyl peptidase IV (EC 3.4.14.5) (DPP IV) (T-cell activation
DE antigen CD26) (TP103) (Adenosine deaminase complexing protein-2)
DE (AABAB).
GN DEPP4 OR ADCP2 OR CD26.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=92329551; PubMed=1352704;
RA Misumi Y., Hayashi Y., Arakawa F., Ikehara Y.;
RT "Molecular cloning and sequence analysis of human dipeptidyl
RT peptidase IV, a serine proteinase on the cell surface.";
RL Biochim. Biophys. Acta 1131:333-336(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=95012454; PubMed=7927537;
RA Abbott C.A., Baker E., Sutherland G.R., McCaughan G.W.;
RT "Genomic organization, exact localization, and tissue expression of
RT the human CD26 (dipeptidyl peptidase IV) gene.";
RL Immunogenetics 40:331-338(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Peripheral blood;
RX MEDLINE=92325476; PubMed=1352530;
RA Tanaka T., Cameron D., Seed B., Torimoto Y., Dang N.H., Kameoka J.,
RA Dahlberg H.N., Schlossman S.F., Morimoto C.;
RT "Cloning and functional expression of the T cell activation antigen
RT CD26.";
RL J. Immunol. 149:481-486(1992).
RN [4]
RP ERRATUM.
RX MEDLINE=93171637; PubMed=8094732;
RA Tanaka T.;
RL J. Immunol. 150:2090-2090(1993).
RN [5]
RP SEQUENCE OF 1-551 FROM N.A.
RC TISSUE=Colon;
RX MEDLINE=92165847; PubMed=1347043;
RA Darmoul D., Lacasa M., Baricault L., Marguet D., Sapin C.,
RA Trottet P., Barbat A.;
RT "Dipeptidyl peptidase IV (CD 26) gene expression in enterocyte-like
RT colon cancer cell lines HT-29 and Caco-2. Cloning of the complete
RT human coding sequence and changes of dipeptidyl peptidase IV mRNA
RT levels during cell differentiation.";
```

```

RL J. Biol. Chem. 267:4824-4833(1992).
RN [6]
RP SEQUENCE OF 545-766 FROM N.A.
RC TISSUE=Colon;
RX MEDLINE=91024044; PubMed=1977364;
RA Darmoul D., Lacasa M., Chantrel I., Swallow D., Trugnan G.;
RT "Isolation of a cDNA probe for the human intestinal
RT dipeptidyl peptidase IV and assignment of the gene locus DEPP4 to
RT chromosome 2.";
RL Ann. Hum. Genet. 54:191-197(1990).
RN [7]
RP SEQUENCE OF 1-31 FROM N.A.
RX MEDLINE=96067599; PubMed=7487939;
RA Boehm S.K., Gum J.R., Jr., Erickson R.H., Hicks J.W., Kim Y.S.;
RT "Isolation of a cDNA probe for the human intestinal
RT dipeptidyl peptidase IV gene promoter. Tissue-specific
RT regulation from a 147A-less GC-rich sequence characteristic of a
RT housekeeping gene promoter.";
RL Biochem. J. 311:835-843(1995).
RN [8]
RP PARTIAL SEQUENCE.
RC TISSUE=Kidney;
RX MEDLINE=93210468; PubMed=8096237;
RA Morrison M.E., Vijayasatradhi S., Engelstein D., Albino A.P.,
RA Houghton A.N.;
RT "A marker for neoplastic progression of human melanocytes is a cell
RT surface ectopeptidase.";
RL J. Exp. Med. 177:1135-1143(1993).
CC -1- FUNCTION: REMOVES N-TERMINAL DIPEPTIDES SEQUENTIALLY FROM
CC POLYPEPTIDES HAVING UNSUBSTITUTED N-TERMINI PROVIDED THAT THE
CC PENULTIMATE RESIDUE IS PROLINE. PLAYS A ROLE IN T CELL ACTIVATION.
CC -1- CATALYTIC ACTIVITY: DIPEPTIDYL-POLYPEPTIDE + H(2)O = DIPEPTIDE +
CC POLYPEPTIDE.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS IN
CC A SOLUBLE FORM.
CC -1- PTM: THE SOLUBLE FORM (SDPP) DERIVES FROM THE MEMBRANE FORM (MDPP)
CC BY PROTEOLYTIC PROCESSING.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S9B; ALSO KNOWN AS THE
CC PROLYL OLIGOPEPTIDASE FAMILY.
CC -1- DATABASE: NAME=PROM; NOTE=CD guide CD26 entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd26.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: U13735; AAB60646.1; -
DR EMBL: U13710; AAB60646.1; JOINED.
DR EMBL: U13711; AAB60646.1; JOINED.
DR EMBL: U13712; AAB60646.1; JOINED.
DR EMBL: U13713; AAB60646.1; JOINED.
DR EMBL: U13714; AAB60646.1; JOINED.
DR EMBL: U13715; AAB60646.1; JOINED.
DR EMBL: U13716; AAB60646.1; JOINED.
DR EMBL: U13717; AAB60646.1; JOINED.
DR EMBL: U13718; AAB60646.1; JOINED.
DR EMBL: U13719; AAB60646.1; JOINED.
DR EMBL: U13720; AAB60646.1; JOINED.
DR EMBL: U13721; AAB60646.1; JOINED.
DR EMBL: U13722; AAB60646.1; JOINED.
DR EMBL: U13723; AAB60646.1; JOINED.
DR EMBL: U13724; AAB60646.1; JOINED.
DR EMBL: U13725; AAB60646.1; JOINED.
DR EMBL: U13726; AAB60646.1; JOINED.
DR EMBL: U13727; AAB60646.1; JOINED.
DR EMBL: U13728; AAB60646.1; JOINED.
DR EMBL: U13729; AAB60646.1; JOINED.
DR EMBL: U13730; AAB60646.1; JOINED.
DR EMBL: U13731; AAB60646.1; JOINED.
```

DR EMBL: U13732; AAB60646.1; JOINED.
 DR EMBL: U13733; AAB60646.1; JOINED.
 DR EMBL: U13734; AAB60646.1; JOINED.
 DR EMBL: M74777; AAB51943.1; -
 DR EMBL: M80536; AAB52308.1; -
 DR EMBL: X60708; CA43118.1; -
 DR EMBL: S79876; AAB35614.1; -
 DR PIR: S24313; CDH026.
 DR MEMOPS: S09.003; -
 DR MIM: 102720; -
 DR InterPro: IPR002469; DEPIV_N_term.
 DR InterPro: IPR000379; Est_11p_choestactsite.
 DR InterPro: IPR001375; Peptidase_S9.
 DR InterPro: IPR002471; Prol_endopep_ser.
 DR Pfam: PF00930; DEPIV_N_term; 1.
 DR Pfam: PF00326; Peptidase_S9; 1.
 DR PROSITE: PS00708; PRO_ENDOPEP_SER; 1.
 DR Hydrolase: Aminopeptidase; Dipeptidase; Serine protease;
 KW Transmembrane; Glycoprotein; Signal-anchor;
 FT CHAIN 1 766
 FT CHAIN 39 766
 FT DOMAIN 1 6
 FT TRANSMEM 7 28
 FT DOMAIN 29 766
 FT ACT_SITE 630 630
 FT ACT_SITE 708 708
 FT ACT_SITE 740 740
 FT CARBOHYD 85 85
 FT CARBOHYD 92 92
 FT CARBOHYD 150 150
 FT CARBOHYD 219 219
 FT CARBOHYD 229 229
 FT CARBOHYD 281 281
 FT CARBOHYD 321 321
 FT CARBOHYD 520 520
 FT CARBOHYD 685 685
 FT CONFLICT 6 6
 FT CONFLICT 7 7
 FT CONFLICT 437 437
 FT CONFLICT 557 557
 FT CONFLICT 663 663
 FT SEQUENCE 766 AA: 88278 MW: 5FBA42C6652D6117 CRC64:
 Query Match 6.9%; Score 87; DB 1; Length 766;
 Best Local Similarity 24.2%; Pred. No. 4.7;
 Matches 48; Conservative 20; Mismatches 66; Indels 64; Gaps 11;
 QY 12 FYGGGDSAGTGTGCGYGNLYS---TGYNTALSS-----TVLFNDGACRSCY 57
 DB 416 YVLSNEYKMGPRGGR---NLKIKQLSDYTKVCLSCGLNPERCQYYSFSKKA---KYY 468
 QY 58 ELRLDNGC---QMLPGSVYVYATNLCPPNYALP-----NDGSGWGN- 96
 DB 469 QLRKSGRLPLVYLHSSVNDKGLRVLEDNSALDKMLQNVMPKSLDFITLNETKFWQM 528
 QY 97 --PPRHFMADEPAFLQIGYRAGIVPVSYRRVPCVKKGCIFFTINGHSYFLVLTNVA 154
 DB 529 ILP--PHFQSKKKYPLLDVY-AG-----PCSQKADIVYFLNWMATY-----LA 568
 QY 155 GPDDVQSVSLKSGSGTGW 172
 DB 569 STENITVASFDGSGCYQ 586
 ID YMV2.CABEL. STANDARD; PRT; 1246 AA.
 AC P34504; P34505; P34506; P90907;
 DT 01-FEB-1994 (rel. 28, Created)
 DT 30-MAY-2000 (rel. 39, Last sequence update)
 DT 01-MAR-2002 (rel. 41, Last annotation update)
 DE Hypothetical 130.6 kDa protein K04H4.2 in chromosome III.

GN K04H4.2.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RC MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,
 RA Crayton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
 RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
 RA Johnson L., Jones M., Kershaw J., Kistler J., Lister N.,
 RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
 RA Sims M., Smaildon N., Smith A., Smith M., Sonnenhammer E., Staden R.,
 RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
 RA Waterston R., Watson A., Wellstock L., Wilkinson-Sproat J.,
 RA Woldman P.;
 RA "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans.";
 RL Nature 368:32-38(1994).
 RN [2]
 RP REVISIONS.
 RA Durbin R.;
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: K04H4.2A (SHOWN HERE) AND
 CC K04H4.2B; ARE PROBABLY PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- SIMILARITY: TO CHITIN-BINDING MOTIFS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed, usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb-sdb.ch).
 CC -----
 DR EMBL: Z27078; CAAB1587.1; -
 DR EMBL: Z27078; CAAB1588.1; -
 DR PIR: S40992; S40992.
 DR PIR: S40993; S40993.
 DR PIR: S40994; S40994.
 DR HSP: P10968; 2CWC.
 DR WormPep: K04H4.2A; CE19967.
 DR WormPep: K04H4.2B; CE19968.
 DR InterPro: IPR002557; Chitin_binding.
 DR InterPro: IPR002899; EB.
 DR SMART: SM00494; ChtBD2; 1.
 DR SMART: SM00289; WRI; 14.
 KW Hypothetical protein; Alternative splicing.
 FT VARSPLIC 166 529
 FT TTPPTPTPTPKLRLSSNTDEVNSLIGASSSSATGCGYN
 ANCTSDDECPPTFKCYOGCCCLAVCPRLATVAKFTCKTOYI
 CRANHECFEGGCCPTTILAVIKSOVLTKDNHEHTETEK
 LIIGCEVDTRVKKCDIILICPEMSECDVGCOPKARC
 GNGMALSLPVHCSLSDCPILASRCCEKCCPPLSEASNDT
 SDSVGETTPVITKEITATKRVKKKDKKISGVSTINKKL
 STQRCDDLTLCPDPTGLSGKCCCKLNHCHDPTGVPTISQ
 SASNHDCPSSSKCTTLNKHFAVCSPGLVGVSTVAVY
 SSECPIGSEVDPRFGTSCRYSLQCPSPYFCNGRQ -> Q
 (IN ISOFORM K04H4.2B).
 DSDTSTNTPSPQETTTKTKKSSSKKRRKKRKKDVP
 LSDPLONDPTIGPGYCGFPHLSNLDVILRADGDECTA
 GLHCTPAINLCCPLILPTIDPKNPKKTKRTRKKQKQDNEM
 EASANFPDSDARFSSISGCGMG -> VG (IN
 ISOFORM K04H4.2B).
 FT VARSPLIC 570 715
 FT LADPLONDPTIGPGYCGFPHLSNLDVILRADGDECTA
 GLHCTPAINLCCPLILPTIDPKNPKKTKRTRKKQKQDNEM
 EASANFPDSDARFSSISGCGMG -> VG (IN
 ISOFORM K04H4.2B).
 FT SEQUENCE 1246 AA: 130610 MW: 4FA1A17D3F9606C4 CRC64:
 Query Match 6.9%; Score 86.5; DB 1; Length 1246;
 Best Local Similarity 23.3%; Pred. No. 8.8;
 Matches 50; Conservative 16; Mismatches 82; Indels 67; Gaps 9;

```

OY 18 ASGTMGCTCGYGNLYSTGYNTALSTLFLPNDGA-ACRSCYELRCNDG-----QMCL 69
DB 824 AGGCYVNGCGCGGYTCSMGLCCAGTSTYKCLDSDAGACIP-SCIGDGGGQVGYVYGG 882
OY 70 PGSVYVATNLCPPNYALPNDGGCMNPPRHEDEMAPEALQGVYRAGIVPYSRYRVC 129
DB 883 SG-YTCTTGNICCPINSCP-N--GGEV-----LGPRTINGLCPRTGY----- 918
OY 130 VKKGIRFTINGHSFNLVLTNVAAGPDQVSIRKSGSTGWMQMSNNQNMQNSYLD 189
DB 919 -----TYOGLNCCSAYCTDGTGLPSVNGVCT-----D 946
OY 190 GOSLSEYVAVSDGRTVTSNNVYPAGMOFGQTEEGG 224
DB 947 GYSLTNGVCCP--ASVYCTDELSTGPGCTGFENG 979

RESULT 14
FLAL_CAMJE STANDARD: PRT: 571 AA.
ID FLAL_CAMJE STANDARD: PRT: 571 AA.
AC P56963: 09PMQO:
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Flagellin A.
GN FLA A OR CJI339C.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
OC Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCCTC 11168;
RX MEDLINE=20150912; PubMed=10688204;
RA Parkhill J., Wren B.W., Mungall K., Kelley J.M., Churcher C.,
RA Basham D., Chillingworth T., Davies R.M., Feltham T., Holtroyd S.,
RA Jagsels K., Kariyeh A.V., Moule S., Pallen M.J., Penn C.W.,
RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
RA Whitehead S., Barrell B.G.;
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT reveals hypervariable sequences."
RU Nature 403:665-668(2000).
CC -1- FUNCTION: FLAGELLIN IS THE SUBUNIT PROTEIN WHICH POLYMERIZES TO
CC FORM THE FILAMENTS OF BACTERIAL FLAGELLA.
CC -1- SUBUNIT: HETEROPOLYMER OF FLA A AND FLA B.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL FLAGELLIN FAMILY.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL FLAGELLIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AL39078: CAB73766.1:
DR InterPro: IPR001028: Flagellin_C.
DR InterPro: IPR001492: Flagellin_N.
DR PRINTS: PR00207: FLAGELLIN.
DR ProDom: PD000316: Flagellin_C: 1.
KW Flagella; Complete proteome.
FT INIT MET 0
FT BY SIMILARITY.
SQ SEQUENCE 571 AA; 58907 MW; 9C63B2A10C1AE863 CRC64;

Query Match 6.8%; Score 85; DB: 1; Length 571;
Best Local Similarity 23.4%; Pred. No. 5.1;
Matches 51; Conservative 26; Mismatches 71; Indels 70; Gaps 10;

OY 2 YSSWOSAHATPYGGDASGTGTCGCGYNTGYTAA-----LSTVL-----FN 48
DB 333 YSS-YSATWSSAGSGFSSGS-GYVSGSKNKTSTGFANALISAASQSLTYVNVASGSGFS 450
OY 49 DGAAAC-----RSCYELRCNDGQWCLPGSVYVATNLCPPNYALPNDGGCMNPPRH 101

```

```

DB 451 SGTSLSEAFATKTTAFGVKDETAATVTLKGAAMV----- 484
OY 102 FDMAPAFLOIGVYRAGIVPYSRYRVPCKKGIPTINGHSFNLVLTNV-----AGPD 158
DB 485 MDIAETALTNDQIRADI-----GSVQNVY--STINNTVYQVWVAESQ 529
OY 159 VOSVIRKSSGTGWMQMSNNQNMQNSYLDGSLSPQ 196
DB 530 IRDVIDFAESANYSKAN-----TLAQSSTYMAQAQNSVQ 563

RESULT 15
CICB_BACTG STANDARD: PRT: 1176 AA.
ID CICB_BACTG STANDARD: PRT: 1176 AA.
AC P56953:
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Pesticidal crystal protein cryIcB (Insecticidal delta-endotoxin
DE cryIcB) (Crystalline entomocidal protoxin) (133 kDa crystal protein).
GN CRYICB OR CRYIC(B).
OS Bacillus thuringiensis (subsp. galleriae).
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=29338;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HD-29;
RX MEDLINE=93236401; PubMed=8476286;
RA Kalman S.S.;
RT "Cloning a novel cryIc-type gene from a strain of Bacillus
RT thuringiensis galleriae."
RT Appl. Environ. Microbiol. 59:1131-1137(1993).
CC -1- FUNCTION: PROMOTES COLICIDEMOTIC LYSIS BY BINDING TO THE MIDGUT
CC EPITHELIAL CELLS OF INSECTS. TOXIC TO SPODOPTERA EXIGUA AND
CC TRITRIPLOUSA NI.
CC -1- DEVELOPMENTAL STAGE: THE CRYSTAL PROTEIN IS PRODUCED DURING
CC SPOROGATION AND IS ACCUMULATED BOTH AS AN INCLUSION AND AS PART
CC OF THE SPORE COAT.
CC -1- MISCELLANEOUS: TOXIC SEGMENT OF THE PROTEIN IS LOCATED IN THE
CC N-TERMINUS.
CC -1- SIMILARITY: BELONGS TO THE DELTA ENDOTOXIN FAMILY.
CC -1- SIMILARITY: BELONGS TO THE DELTA ENDOTOXIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M97880: -; NOT_ANNOTATED_CDS.
DR HSSP: P02965: ICIV.
DR InterPro: IPR001178: Endotoxin.
DR Pfam: PF00555: endotoxin; 1.
KW Toxin; Sporulation.
SQ SEQUENCE 1176 AA; 132867 MW; 108233494C2AC018 CRC64;

Query Match 6.7%; Score 84.5; DB: 1; Length 1176;
Best Local Similarity 22.5%; Pred. No. 12;
Matches 66; Conservative 30; Mismatches 98; Indels 99; Gaps 17;

OY 2 YSSWOSAHATPYGGD--ASGTGCTGCGYNTGYTAA-----TNTAALSTVLFNDGACRSC 56
DB 310 FTDWFSVGRNRYWGHRYTSNRIGG-----GNITSPIGRANDEPPRSPFTFN-GVYFRTL 364
OY 57 -----YELRCNDGQWCLP-GSVYVATNLCPPNYALPNDGGCMNPPR 99
DB 365 SNPTFRPLQGWPPAPFNLKRGVEVERSTPLNSTYRGRTVDSLTELPEPDNS--VPPR 422
OY 100 PHFD--MAPEALQIGVYRAG-----IVPYSRYRVCVK----- 131

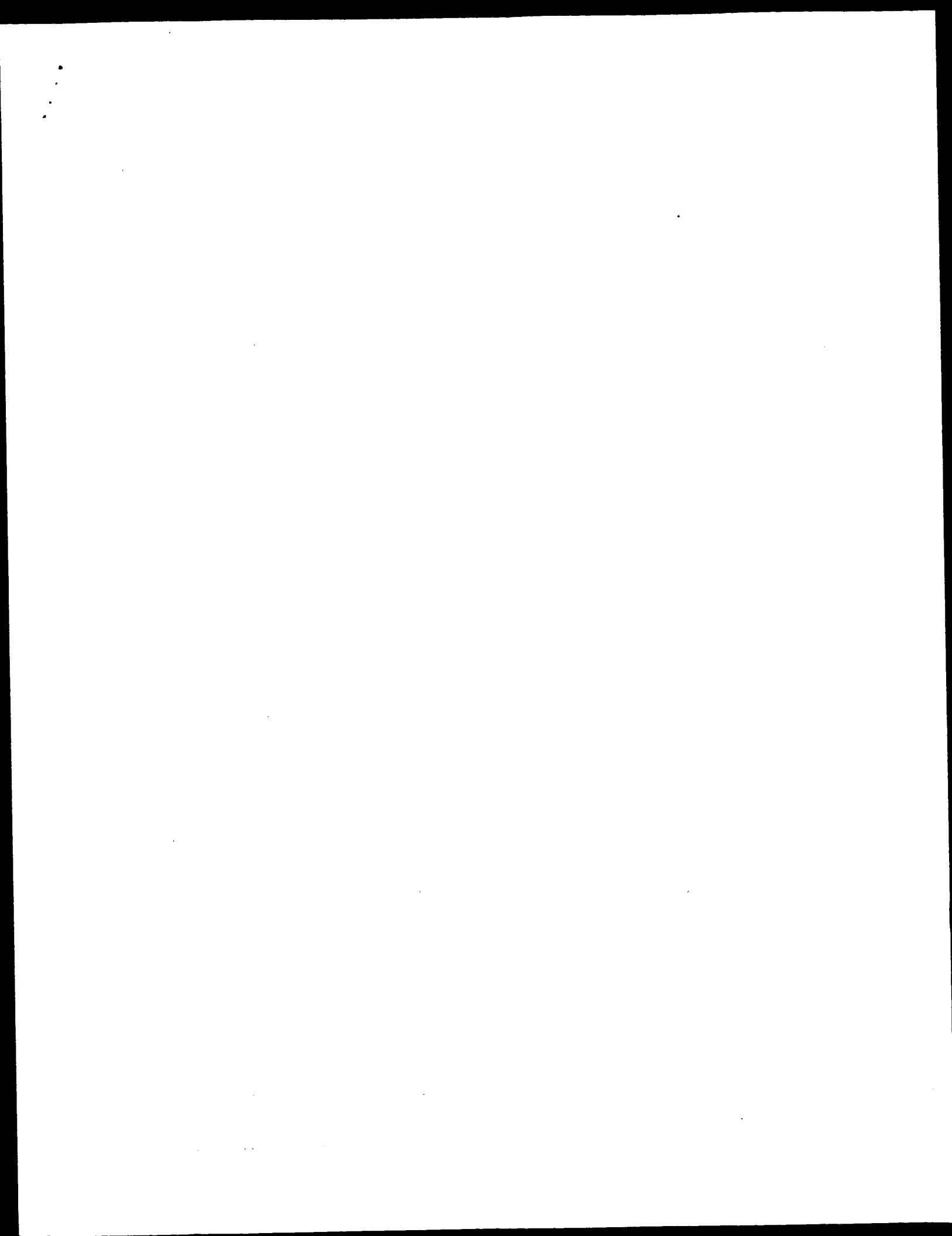
```

```

Db 423 EGYSHRLCHATEVO---RSGTPELTTPGVFSWTHRSATDRNIIYPDVINQIPLVKAFNL 478
OY 132 -----KG--IREPTINGH-----SYENLVL-----VTNAGPBGDVOSVSIK 165
Db 479 TSGTSVVRGPGFTGCDIIRTNVNGSVLSMSLNFSTTLORYRVRVRYAASQTMVMSVTVG 538
OY 166 GSSTGMOPMRNGQNMQNSYLDGOSLSF-----QVAVSDGRTVT-SNNV 210
Db 539 GSTTG-----NOCFPSTMSANGALTSQSFRFAEPYGISASGASISISNNV 587

```

Search completed: October 11, 2002, 14:57:35
 Job time : 7.83764 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 11, 2002, 14:54:05 : Search time 8.83985 Seconds
(without alignments)
624,466 Million cell updates/sec

Title: US-09-896-301-6

Perfect score: 1257
Sequence: 1 DYSSWQSAHATFYGGDASG.....SNNVPAQMGQFQTEFGQF 226

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Issued Patents_AA:*

- 1: /cgn2_6/ptodata/2/1aa/5a_COMB.pep:*
- 2: /cgn2_6/ptodata/2/1aa/5b_COMB.pep:*
- 3: /cgn2_6/ptodata/2/1aa/6a_COMB.pep:*
- 4: /cgn2_6/ptodata/2/1aa/6b_COMB.pep:*
- 5: /cgn2_6/ptodata/2/1aa/pctus_COMB.pep:*
- 6: /cgn2_6/ptodata/2/1aa/backfillsl.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1257	100.0	226	2 US-08-440-517A-6	Sequence 6, Appli
2	1257	100.0	226	4 US-09-092-160-6	Sequence 6, Appli
3	1021.5	81.3	227	4 US-09-092-160-7	Sequence 7, Appli
4	959.5	76.3	228	4 US-08-440-517A-2	Sequence 2, Appli
5	959.5	76.3	228	4 US-08-440-517A-2	Sequence 2, Appli
6	863.5	68.7	225	2 US-08-440-517A-5	Sequence 5, Appli
7	863.5	68.7	225	2 US-08-440-517A-5	Sequence 5, Appli
8	791.5	63.0	225	2 US-08-845-539-2	Sequence 2, Appli
9	787	62.6	222	2 US-08-440-517A-3	Sequence 3, Appli
10	787	62.6	222	4 US-09-092-160-3	Sequence 3, Appli
11	783.5	62.3	225	4 US-09-362-642-2	Sequence 3, Appli
12	742	59.0	227	2 US-08-440-517A-4	Sequence 4, Appli
13	742	59.0	227	2 US-08-440-517A-4	Sequence 4, Appli
14	726.5	57.8	179	4 US-08-845-539-6	Sequence 6, Appli
15	726.5	57.8	179	4 US-08-845-539-6	Sequence 6, Appli
16	682.5	54.3	167	2 US-08-362-642-6	Sequence 6, Appli
17	682.5	54.3	167	2 US-08-362-642-6	Sequence 6, Appli
18	551.5	50.0	261	1 US-09-362-642-4	Sequence 4, Appli
19	251.5	20.0	261	1 US-07-971-096-2	Sequence 2, Appli
20	225	17.9	246	4 US-08-175-096-2	Sequence 2, Appli
21	219	17.4	272	4 US-08-441-507-21	Sequence 15, Appli
22	213.5	17.0	263	1 US-07-971-096-4	Sequence 4, Appli
23	213.5	17.0	263	1 US-08-175-096-4	Sequence 4, Appli
24	213.5	17.0	263	4 US-08-413-974-6	Sequence 6, Appli
25	213.5	17.0	263	4 US-08-433-418-6	Sequence 6, Appli
26	213.5	17.0	263	4 US-08-433-288-6	Sequence 6, Appli
27	213.5	17.0	263	4 US-08-174-739A-6	Sequence 6, Appli

28	202	16.1	245	4 US-08-441-507-24	Sequence 24, Appli
29	169	13.4	200	4 US-08-441-507-4	Sequence 4, Appli
30	168	13.4	197	4 US-08-441-507-5	Sequence 5, Appli
31	125.5	10.0	122	4 US-08-441-507-23	Sequence 23, Appli
32	109.5	8.7	145	4 US-08-413-974-4	Sequence 4, Appli
33	109.5	8.7	145	4 US-08-434-418-4	Sequence 4, Appli
34	109.5	8.7	145	4 US-08-433-288-4	Sequence 4, Appli
35	109.5	8.7	145	4 US-08-174-739A-4	Sequence 4, Appli
36	107.5	8.6	145	4 US-08-441-507-2	Sequence 2, Appli
37	106.5	8.5	145	4 US-08-441-507-16	Sequence 16, Appli
38	102	8.1	138	4 US-08-144-121-3	Sequence 3, Appli
39	89	7.1	1147	1 US-08-144-121-3	Sequence 3, Appli
40	89	7.1	1147	2 US-08-735-893-3	Sequence 3, Appli
41	89	7.1	1165	1 US-08-144-121-2	Sequence 2, Appli
42	87	7.1	1165	2 US-08-735-893-3	Sequence 2, Appli
43	87	6.9	593	5 PCT-US93-07923-11	Sequence 11, Appli
44	87	6.9	755	5 PCT-US93-07923-3	Sequence 3, Appli
45	87	6.9	759	5 PCT-US93-07923-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-08-440-517A-6

Sequence 6, Application US/08440517A
Patent No. 5959082

GENERAL INFORMATION:

APPLICANT: COSGROVE, DANIEL J.;

APPLICANT: GUILTINAN, MARK;

APPLICANT: SCHERBAN, TATYANA;

APPLICANT: SHI, JUN

TITLE OF INVENTION: PURIFIED EXPANSIN PROTEINS

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESSES:

ADDRESSEE: INTELLECTUAL PROPERTY OFFICE, THE

STREET: 113 TECHNOLOGY CENTER

CITY: UNIVERSITY PARK

STATE: PENNSYLVANIA

COUNTRY: UNITED STATES OF AMERICA

ZIP: 16802-7000

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK

COMPUTER: NEC 286

OPERATING SYSTEM: DOS

SOFTWARE: WORDPERFECT 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/440,517A

FILING DATE:

CLASSIFICATION: 530

INFORMATION FOR SEQ. ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 226

TYPE: AMINO ACID

TOPOLOGY: UNKNOWN

US-08-440-517A-6

Query Match

Best Local Similarity 100.0%; Score 1257; DB 2; Length 226;

Matches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	DYSSWQSAHATFYGGDASGTCGYNLVSTGYNTAALSTVLPNDGACRCSEYLR	60
DB	1	DYSSWQSAHATFYGGDASGTCGYNLVSTGYNTAALSTVLPNDGACRCSEYLR	60
QY	61	CDNDGQWCLPGSVTVATNLCPNTVALPNDGQWGNPPRPHDMAEPFLQIGVRRACIV	120
DB	61	CDNDGQWCLPGSVTVATNLCPNTVALPNDGQWGNPPRPHDMAEPFLQIGVRRACIV	120
QY	121	PVSYRRVPCVKKGGIFETNGHSYFLVYTVNAPRGDQVSITGSGTQWQPSRRNQO	180
DB	121	PVSYRRVPCVKKGGIFETNGHSYFLVYTVNAPRGDQVSITGSGTQWQPSRRNQO	180

QY 181 NMOSNSYLDGOSLSFOVAVSDGRTVTSNNVYPAGMOFGQTFFEGGOF 226
 |||
 DB 181 NMOSNSYLDGOSLSFOVAVSDGRTVTSNNVYPAGMOFGQTFFEGGOF 226

RESULT 2
 US-09-092-160-6
 ; Sequence 6, Application US/09092160C

; Patent No. 6255466
 ; GENERAL INFORMATION:
 ; APPLICANT: Cosgrove, Daniel J
 ; APPLICANT: McQueen-Mason, Simon
 ; APPLICANT: Guillinan, Mark J
 ; APPLICANT: Shcherban, Tatyana
 ; APPLICANT: Shi, Jun
 ; TITLE OF INVENTION: PURIFIED EXPANSIN PROTEINS
 ; FILE REFERENCE: 1194/JC114US3
 ; CURRENT APPLICATION NUMBER: US/09/092,160C
 ; CURRENT FILING DATE: 1998-06-05
 ; EARLIER APPLICATION NUMBER: 08/440,517
 ; EARLIER FILING DATE: 1995-05-12
 ; EARLIER APPLICATION NUMBER: 08/242,090
 ; EARLIER FILING DATE: 1994-05-12
 ; EARLIER APPLICATION NUMBER: 08/060,944
 ; EARLIER FILING DATE: 1993-05-12
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 6
 ; LENGTH: 226
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Arabidopsis
 ; OTHER INFORMATION: expansin
 US-09-092-160-6

Query Match 100.0%; Score 1257; DB 4; Length 226;
 Best Local Similarity 100.0%; Pred. No. 6.7e-112;
 Matches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYSWOSAHATFYGGDASGTGTCGYGNLYSTGYTNTALSTVLFNDGAACRSCYEEL 60
 |||
 DB 1 DYSWOSAHATFYGGDASGTGTCGYGNLYSTGYTNTALSTVLFNDGAACRSCYEEL 60
 QY 61 CDNDGQWCLPGSVYVATNLCPPNYALPNDGQWCPNPPHEDMAEPALQIGVYRAGI 120
 |||
 DB 61 CDNDGQWCLPGSVYVATNLCPPNYALPNDGQWCPNPPHEDMAEPALQIGVYRAGI 120
 QY 121 PVSYRRVPCYKKGIRFTINGHSYFNLYVTNVAAGPDVQSVSIKGSSTGQPMSSRMWG 180
 |||
 DB 121 PVSYRRVPCYKKGIRFTINGHSYFNLYVTNVAAGPDVQSVSIKGSSTGQPMSSRMWG 180
 QY 181 NMOSNSYLDGOSLSFOVAVSDGRTVTSNNVYPAGMOFGQTFFEGGOF 226
 |||
 DB 181 NMOSNSYLDGOSLSFOVAVSDGRTVTSNNVYPAGMOFGQTFFEGGOF 226

RESULT 3
 US-09-092-160-7
 ; Sequence 7, Application US/09092160C
 ; Patent No. 6255466
 ; GENERAL INFORMATION:
 ; APPLICANT: Cosgrove, Daniel J
 ; APPLICANT: McQueen-Mason, Simon
 ; APPLICANT: Guillinan, Mark J
 ; APPLICANT: Shcherban, Tatyana
 ; APPLICANT: Shi, Jun
 ; TITLE OF INVENTION: PURIFIED EXPANSIN PROTEINS
 ; FILE REFERENCE: 1194/JC114US3
 ; CURRENT APPLICATION NUMBER: US/09/092,160C
 ; CURRENT FILING DATE: 1998-06-05
 ; EARLIER APPLICATION NUMBER: 08/440,517

; EARLIER FILING DATE: 1995-05-12
 ; EARLIER APPLICATION NUMBER: 08/242,090
 ; EARLIER FILING DATE: 1994-05-12
 ; EARLIER APPLICATION NUMBER: 08/060,944
 ; EARLIER FILING DATE: 1993-05-12
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 7
 ; LENGTH: 227
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: cucumber
 ; OTHER INFORMATION: expansin
 US-09-092-160-7

Query Match 81.3%; Score 1021.5; DB 4; Length 227;
 Best Local Similarity 78.0%; Pred. No. 1.5e-89;
 Matches 177; Conservative 21; Mismatches 28; Indels 1; Gaps 1;

QY 1 DYSWOSAHATFYGGDASGTGTCGYGNLYSTGY-TNTALSTVLFNDGAACRSCYEEL 59
 |||
 DB 1 DYSWOSAHATFYGGDASGTGTCGYGNLYSTGY-TNTALSTVLFNDGAACRSCYEEL 60
 QY 60 RCDNDGQWCLPGSVYVATNLCPPNYALPNDGQWCPNPPHEDMAEPALQIGVYRAGI 119
 |||
 DB 61 TCTNDPKWCLPGIRVYVATNLCPPNYALPNDGQWCPNPPHEDMAEPALQIGVYRAGI 120
 QY 120 PVSYRRVPCYKKGIRFTINGHSYFNLYVTNVAAGPDVQSVSIKGSSTGQPMSSRMWG 179
 |||
 DB 121 PVSYRRVPCYKKGIRFTINGHSYFNLYVTNVAAGPDVQSVSIKGSSTGQPMSSRMWG 180
 QY 180 NMOSNSYLDGOSLSFOVAVSDGRTVTSNNVYPAGMOFGQTFFEGGOF 226
 |||
 DB 181 NMOSNSYLDGOSLSFOVAVSDGRTVTSNNVYPAGMOFGQTFFEGGOF 227

RESULT 4
 US-08-440-517A-2
 ; Sequence 2, Application US/08440517A
 ; Patent No. 5959082

; GENERAL INFORMATION:
 ; APPLICANT: COSGROVE, DANIEL J.;
 ; APPLICANT: GUILLINAN, MARK;
 ; APPLICANT: SCHERBAN, TATYANA;
 ; APPLICANT: SHI, JUN
 ; TITLE OF INVENTION: PURIFIED EXPANSIN PROTEINS
 ; NUMBER OF SEQUENCES: 6
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: INTELLECTUAL PROPERTY OFFICE, THE
 ; ADDRESSEE: PENNSYLVANIA STATE UNIVERSITY
 ; STREET: 113 TECHNOLOGY CENTER
 ; CITY: UNIVERSITY PARK
 ; STATE: PENNSYLVANIA
 ; COUNTRY: UNITED STATES OF AMERICA
 ; ZIP: 16802-7000
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: FLOPPY DISK
 ; COMPUTER: NEC 286
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: WORDPERFECT 5.1
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/440,517A
 ; FILING DATE:
 ; CLASSIFICATION: 530
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 228
 ; TYPE: AMINO ACID
 ; TOPOLOGY: UNKNOWN
 US-08-440-517A-2

Query Match 76.3%; Score 959.5; DB 2; Length 228;

Best Local Similarity 75.0%; Pred. No. 1,2e-83;
Matches 165; Conservative 23; Mismatches 31; Indels 1; Gaps 1;

QY 5 MOSAHATFYGGDASGTMGTCGYGNYSTGY-TNTALSTVLFNDGAACRSCELRCDN 63
1 :|||||
Db 5 WYNAHATFYGGDASGTMGACGYGNYSGYGTNTALSTALFNNGLSGCACFEIRCON 64
1 :|||||
QY 64 DQOMCLPGSVTVATNLCPPNYALPNDGCMCNPPRPHDMAEPFLQIGYRAGIYPV 123
1 :|||||
Db 65 DQKMLPGSVTVATNFCPCPNALPNNAGCMCNPPQOIFDLSQPVFORIAQYRAGIY 124
1 :|||||
QY 124 YRRVPCVKKGIRFTINGSHSYFNLYLVTVNAGPDVQSVSIRKSGSTGQPMRNGQNM 183
1 :|||||
Db 125 YRRVPCVKKGIRFTINGSHSYFNLYLVTVNAGPDVQSVSIRKSGSTGQPMRNGQNM 184
1 :|||||
QY 184 SNSYLDGSLSFQVAVSDGRTVTSNNVVPAGQFGOTFEG 223
1 :|||||
Db 185 SNSYLDGSLSFQVAVSDGRTVTSNNVVPAGQFGOTFEG 224
1 :|||||

RESULT 5

US-09-092-160-2
Sequence 2, Application US/09092160C
Patent No. 6255466
GENERAL INFORMATION:
APPLICANT: Cosgrove, Daniel J
APPLICANT: McQueen-Mason, Simon
APPLICANT: Guiltinan, Mark J
APPLICANT: Shepherdan, Tatyana
APPLICANT: Shi, Jun
TITLE OF INVENTION: PURIFIED EXPANSIN PROTEINS
FILE REFERENCE: 1194/1C114U53
CURRENT APPLICATION NUMBER: US/09/092,160C
CURRENT FILING DATE: 1998-06-05
EARLIER APPLICATION NUMBER: 08/440,517
EARLIER FILING DATE: 1995-05-12
EARLIER APPLICATION NUMBER: 08/242,090
EARLIER FILING DATE: 1994-05-12
EARLIER APPLICATION NUMBER: 08/060,944
EARLIER FILING DATE: 1993-05-12
NUMBER OF SEQ ID NOS: 7
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2
LENGTH: 228
TYPE: PRP
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: rice expansin
NAME/KEY: UNSURE
LOCATION: 211
OTHER INFORMATION: Xaa is unknown or other.
US-09-092-160-2

Query Match 76.3%; Score 959.5; DB 4; Length 228;
Best Local Similarity 75.0%; Pred. No. 1,2e-83;

Matches 165; Conservative 23; Mismatches 31; Indels 1; Gaps 1;

QY 5 MOSAHATFYGGDASGTMGTCGYGNYSTGY-TNTALSTVLFNDGAACRSCELRCDN 63
1 :|||||
Db 5 WYNAHATFYGGDASGTMGACGYGNYSGYGTNTALSTALFNNGLSGCACFEIRCON 64
1 :|||||
QY 64 DQOMCLPGSVTVATNLCPPNYALPNDGCMCNPPRPHDMAEPFLQIGYRAGIYPV 123
1 :|||||
Db 65 DQKMLPGSVTVATNFCPCPNALPNNAGCMCNPPQOIFDLSQPVFORIAQYRAGIY 124
1 :|||||
QY 124 YRRVPCVKKGIRFTINGSHSYFNLYLVTVNAGPDVQSVSIRKSGSTGQPMRNGQNM 183
1 :|||||
Db 125 YRRVPCVKKGIRFTINGSHSYFNLYLVTVNAGPDVQSVSIRKSGSTGQPMRNGQNM 184
1 :|||||
QY 184 SNSYLDGSLSFQVAVSDGRTVTSNNVVPAGQFGOTFEG 223
1 :|||||
Db 185 SNSYLDGSLSFQVAVSDGRTVTSNNVVPAGQFGOTFEG 224
1 :|||||

RESULT 6

US-08-440-517A-5
Sequence 5, Application US/08440517A
Patent No. 5959082
GENERAL INFORMATION:
APPLICANT: COSGROVE, DANIEL J.;
APPLICANT: GUILTINAN, MARK;
APPLICANT: SHEPHERDAN, TATYANA;
APPLICANT: SHI, JUN
TITLE OF INVENTION: PURIFIED EXPANSIN PROTEINS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: INTELLECTUAL PROPERTY OFFICE, THE
ADDRESSEE: PENNSYLVANIA STATE UNIVERSITY
STREET: 113 TECHNOLOGY CENTER
CITY: UNIVERSITY PARK
STATE: PENNSYLVANIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 16802-7000
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: NEC 286
OPERATING SYSTEM: DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/440,517A
FILING DATE:
CLASSIFICATION: 530
INFORMATION FOR SEQ ID NO: 5
SEQUENCE CHARACTERISTICS:
LENGTH: 225
TYPE: AMINO ACID
TOPOLOGY: UNKNOWN
US-08-440-517A-5

Query Match

Best Local Similarity 68.7%; Score 863.5; DB 2; Length 225;
Matches 158; Conservative 24; Mismatches 41; Indels 5; Gaps 4;

QY 1 DYSSMOAHATFYGGDASGTMGTCGYGNYSTGY-TNTALSTVLFNDGAACRSCELR 59
1 :|||||
Db 1 DNGMGRGHATFYGGDASGTMGACGYGNYSGYGTNTALSTALFNNGLSGCACFEIR 60
1 :|||||
QY 60 RCDNDGOMCLPGSVTVATNLCPPNYALPNDGCMCNPPRPHDMAEPFLQIGYRAGI 119
1 :|||||
Db 61 TCEDDPEMCLPGSVTVATNLCPPNYALPNDGCMCNPPRPHDMAEPFLQIGYRAGI 117
1 :|||||
QY 120 VPSYRVRPCVKKGIRFTINGSHSYFNLYLVTVNAGPDVQSVSIRKSGSTG-WQPMRNM 178
1 :|||||
Db 118 VPSYRVRPCVKKGIRFTINGSHSYFNLYLVTVNAGPDVQSVSIRKSGSTG-WQPMRNM 177
1 :|||||
QY 179 GQNMQSNLYLDGSLSFQVAVSDGRTVTSNNVVPAGQFGOTFEG 226
1 :|||||
Db 178 GQNMQSNLYLDGSLSFQVAVSDGRTVTSNNVVPAGQFGOTFEG 225
1 :|||||

RESULT 7

US-09-092-160-5
Sequence 5, Application US/09092160C
Patent No. 6255466
GENERAL INFORMATION:
APPLICANT: Cosgrove, Daniel J
APPLICANT: McQueen-Mason, Simon
APPLICANT: Guiltinan, Mark J
APPLICANT: Shepherdan, Tatyana
APPLICANT: Shi, Jun
TITLE OF INVENTION: PURIFIED EXPANSIN PROTEINS
FILE REFERENCE: 1194/1C114U53
CURRENT APPLICATION NUMBER: US/09/092,160C
CURRENT FILING DATE: 1998-06-05
EARLIER APPLICATION NUMBER: 08/440,517

```

; EARLIER FILING DATE: 1995-05-12
; EARLIER APPLICATION NUMBER: 08/242,090
; EARLIER FILING DATE: 1994-05-12
; EARLIER APPLICATION NUMBER: 08/060,944
; EARLIER FILING DATE: 1993-05-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Arabidopsis
; OTHER INFORMATION: expansin
US-09-092-160-5

```

```

Query Match          68.7%; Score 863.5; DB 4; Length 225;
Best Local Similarity 69.3%; Pred. No. 1.5e-74;
Matches 158; Conservative 24; Mismatches 41; Indels 5; Gaps 4;

```

```

QY 1 DYSMSQAHATFYGGDASGTMGTCGYGNLYSTGY-TNTAALSTVLFNDGACRSCYEL 59
   | : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 DNGGMRGATFYGGADSGTMGACGYGNLHSGYGLQTALSTALFNSGKCGACFEL 60
   | : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 60 RCDNDGQWCLPGSVTVATNLCPNYALPNDGCGWCPNPPHFDMAEPALQIGYRAGI 119
   | : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 TCEDDDEMCIPGSI-IVRYNLA--NFALANDGWCNPELKHFDIAEPALQIAYRAGI 117
   | : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 120 VPSYRVRVPCVKKGIRFTINGHSYENLVLTNVAGPDVQSVSIKGSSTG-WQPMRNM 178
   | : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 118 VPAFRVRVPCVKKGIRFTINGHSYENLVLTNVAGPDVQSVSIKGSSTG-WQPMRNM 177
   | : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 179 GQWQSNSTYLDGQSLFQVAVSDGRTVTSNNVPPGQGFQTEGGQF 226
   | : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 178 GQWQSNSTYLDGQSLFQVAVSDGRTVTSNNVPPGQGFQTEGGQF 225
   | : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

RESULT 8
US-08-845-539-2
; Sequence 2, Application US/08845539
; Patent No. 5929303
; GENERAL INFORMATION:
; APPLICANT: Bennett, Alan B.
; TITLE OF INVENTION: Fruit-specific and Ripening-Regulation
; TITLE OF INVENTION: Expansin Genes to Control Fruit Texture and Softening
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/845,539
; FILING DATE: 25-Apr-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-078200US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 225 amino acids

```

```

; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-845-539-2

```

```

Query Match          63.0%; Score 791.5; DB 2; Length 225;
Best Local Similarity 71.1%; Pred. No. 1.1e-67;
Matches 138; Conservative 24; Mismatches 27; Indels 5; Gaps 3;

```

```

QY 4 SWQSAHATFYGGDASGTMGTCGYGNLYSTGY-TNTAALSTVLFNDGACRSCYELRCD 62
   | : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 32 SWEIHAATFYGGDASGTMGACGYGNLYSGYVNTAALSTALFNSGKCGACFELKCI 91
   | : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 63 N--DGQWCLPG--SVTVATNLCPNYALPNDGCGWCPNPPHFDMAEPALQIGYRAGI 118
   | : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 92 NTPNWKCLPGNPILITATNFCPPNYALPNDGCGWCPNPPHFDMAEPALQIGYRAGI 151
   | : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 119 IVPYSYRVRVPCVKKGIRFTINGHSYENLVLTNVAGPDVQSVSIKGSSTGQPMRNM 178
   | : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 152 IVPYTRIRIPCKKGGIRFTINGHSYENLVLTNVAGADILKVKVGTQTNMIPLSRNM 211
   | : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 179 GQWQSNSTYLDGQSL 192
   | : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 212 GQWQSNSTYLDGQSL 225
   | : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

RESULT 9
US-08-440-517A-3
; Sequence 3, Application US/08440517A
; Patent No. 5959082
; GENERAL INFORMATION:
; APPLICANT: COSGROVE, DANIEL J.;
; APPLICANT: GUILTINAN, MARK;
; APPLICANT: SCHERBAN, TATYANA;
; APPLICANT: SHI, JUN
; TITLE OF INVENTION: PURIFIED EXPANSIN PROTEINS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INTELLECTUAL PROPERTY OFFICE, THE
; ADDRESS: PENNSYLVANIA STATE UNIVERSITY
; STREET: 113 TECHNOLOGY CENTER
; CITY: UNIVERSITY PARK
; STATE: PENNSYLVANIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 16802-7000
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: NEC 286
; OPERATING SYSTEM: DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/440,517A
; FILING DATE:
; CLASSIFICATION: 530
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 222
; TYPE: AMINO ACID
; TOPOLOGY: UNKNOWN
US-08-440-517A-3

```

```

Query Match          62.6%; Score 787; DB 2; Length 222;
Best Local Similarity 65.2%; Pred. No. 2.8e-67;
Matches 146; Conservative 21; Mismatches 47; Indels 10; Gaps 3;

```

```

QY 5 WQSAHATFYGGDASGTMGTCGYGNLYSTGY-TNTAALSTVLFNDGACRSCYELRCDN 63
   | : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5 WINHAATFYXXGDXXTMGACGYGNLYSGYGLTALSTALFDGLSGACXELMCVN 64
   | : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 64 DGQWCLPG-SVTVATNLCPNYALPNDGCGWCPNPPHFDMAEPALQIGYRAGIYV 122
   | : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 65 DPWCIKRSIVTVATNCP-----GGACDPNHHFDLSOPLYEKIALYKSSIIIV 116
   | : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

QY 123 SYRRVPCVKKGGIRFTINGHSYFNLVLTNNVAGPDVQSVISKSGSTGQPMRNMGM 182
 Db 117 MYRRVRCRSGIRFTINGHSYFNLVLTNNVAGDVHVSMSKSRFTKQMLSMNMGM 176
 QY 183 QSNSTLDQGSLSFOVAASDGRVTSSNNVVPAGMFGOTFECCGF 226
 Db 177 QSNSTLDQGSLSFVYTTSDRRSVSFNVAPPTWSFGOTYTGQGF 220

RESULT 10

US-09-092-160-3
 Sequence 3, Application US/09092160C
 Patent No. 6253466

GENERAL INFORMATION:
 APPLICANT: Cosgrove, Daniel J.
 APPLICANT: McQueen-Mason, Simon
 APPLICANT: Gullitnan, Mark J.
 APPLICANT: Shcherban, Tatyana
 APPLICANT: Shi, Jun
 TITLE OF INVENTION: PURIFIED EXPANSIN PROTEINS
 FILE REFERENCE: 1194/1C114US3
 CURRENT APPLICATION NUMBER: US/09/092,160C
 CURRENT FILING DATE: 1998-06-05
 EARLIER APPLICATION NUMBER: 08/440,517
 EARLIER FILING DATE: 1993-05-12
 EARLIER APPLICATION NUMBER: 08/242,090
 EARLIER FILING DATE: 1994-05-12
 EARLIER APPLICATION NUMBER: 08/060,944
 EARLIER FILING DATE: 1993-05-12
 NUMBER OF SEQ ID NOS: 7
 SOFTWARE: Patentn Ver. 2.1
 SEQ ID NO 3
 LENGTH: 222
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 NAME/KEY: UNSURE
 LOCATION: (14)..(58)
 OTHER INFORMATION: Xaa is unknown or other.
 US-09-092-160-3

Query Match

62.6%; Score 787; DB 4; Length 222;
 Best Local Similarity 65.2%; Pred. No. 2,8e-67;

Matches 146; Conservative 21; Mismatches 47; Indels 10; Gaps 3;

QY 5 WQSHATFYGGDASGTMGTCGYNLYSTGY-TNTAALSTVLFNDGACRSCYELKCDN 63
 Db 5 WINHATFYXXGDXXTMGAGCGNLYSGYGLTALSTALFDQGLSCGACXELKCVN 64
 QY 64 DGQWCLPG-SVTVTATNLCPNYALPNDGCMCPRPHEFDMAPFLQIGVYRAGIIPV 122
 Db 65 DPQWCLPG-SVTVTATNLCPNYALPNDGCMCPRPHEFDMAPFLQIGVYRAGIIPV 116
 QY 123 SYRRVPCVKKGGIRFTINGHSYFNLVLTNNVAGPDVQSVISKSGSTGQPMRNMGM 182
 Db 117 MYRRVRCRSGIRFTINGHSYFNLVLTNNVAGDVHVSMSKSRFTKQMLSMNMGM 176
 QY 183 QSNSTLDQGSLSFOVAASDGRVTSSNNVVPAGMFGOTFECCGF 226
 Db 177 QSNSTLDQGSLSFVYTTSDRRSVSFNVAPPTWSFGOTYTGQGF 220

RESULT 11

US-09-362-642-2
 Sequence 2, Application US/09362642
 Patent No. 6350935

GENERAL INFORMATION:
 APPLICANT: Bennett, Alan B.
 APPLICANT: Rose, Jocelyn K.C.
 APPLICANT: The Regents of the University of California
 TITLE OF INVENTION: Fruit-Specific and Ripening Regulation Expansin Genes

TITLE OF INVENTION: to Control Fruit Texture and Softening
 FILE REFERENCE: 023070-078210US
 CURRENT APPLICATION NUMBER: US/09/362,642
 CURRENT FILING DATE: 1999-07-27
 NUMBER OF SEQ ID NOS: 8
 SOFTWARE: Patentn Ver. 2.1
 SEQ ID NO 2
 LENGTH: 225
 TYPE: PRT
 ORGANISM: Lycopersicon esculentum cv. T5
 US-09-362-642-2

Query Match

62.3%; Score 783.5; DB 4; Length 225;
 Best Local Similarity 70.6%; Pred. No. 6,1e-67;

Matches 137; Conservative 24; Mismatches 28; Indels 5; Gaps 3;

QY 4 SMOAHATFYGGDASGTMGTCGYNLYSTGY-TNTAALSTVLFNDGACRSCYELKCDN 62
 Db 32 SWEHATFYGGDASGTMGAGCGNLYSGYGVNTALSTALFNNGLSCGACFELKCT 91
 QY 63 N--DGQWCLPG-SVTVTATNLCPNYALPNDGCMCPRPHEFDMAPFLQIGVYRAG 118
 Db 92 NTPMWMCLPGNPSILITATNFCPPNYALPNDGCMCPRPHEFDMAPFLQIGVYRAG 151
 QY 119 IVPVSYRRVPCVKKGGIRFTINGHSYFNLVLTNNVAGPDVQSVISKSGSTGQPMRNMGM 178
 Db 152 IVPVSYRRVPCVKKGGIRFTINGHSYFNLVLTNNVAGDVHVSMSKSRFTKQMLSMNMGM 211
 QY 179 GQWQSNSTLDQGS 192
 Db 212 GQWQSNSTLDQGS 225

RESULT 12

US-08-440-517A-4
 Sequence 4, Application US/08440517A
 Patent No. 5959082

GENERAL INFORMATION:
 APPLICANT: COSGROVE, DANIEL J.;
 APPLICANT: GULLITNAN, MARK;
 APPLICANT: SCHERBAN, TATYANA;
 APPLICANT: SHI, JUN
 TITLE OF INVENTION: PURIFIED EXPANSIN PROTEINS
 NUMBER OF SEQUENCES: 6
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: INTELLECTUAL PROPERTY OFFICE, THE
 ADDRESS: PENNSYLVANIA STATE UNIVERSITY
 STREET: 113 TECHNOLOGY CENTER
 CITY: UNIVERSITY PARK
 STATE: PENNSYLVANIA
 COUNTRY: UNITED STATES OF AMERICA
 ZIP: 16802-7000
 COMPUTER READABLE FORM:
 MEDIUM TYPE: FLOPPY DISK
 COMPUTER: NEC 286
 OPERATING SYSTEM: DOS
 SOFTWARE: WORDPERFECT 5.1
 CURRENT APPLICATION DATA: 5.1
 APPLICATION NUMBER: US/08/440,517A
 FILING DATE:
 CLASSIFICATION: 530
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 227
 TYPE: AMINO ACID
 TOPOLOGY: UNKNOWN
 US-08-440-517A-4

Query Match

59.0%; Score 742; DB 2; Length 227;
 Best Local Similarity 64.0%; Pred. No. 5,4e-63;

Matches 137; Conservative 24; Mismatches 49; Indels 4; Gaps 3;

QY 6 QSHATFYGGDASGTMGTCGYNLYSTGY-TNTAALSTVLFNDGACRSCYELKCDN 63

Sun Oct 13 14:19:08 2002

us-09-896-301-6.rai

Page 7

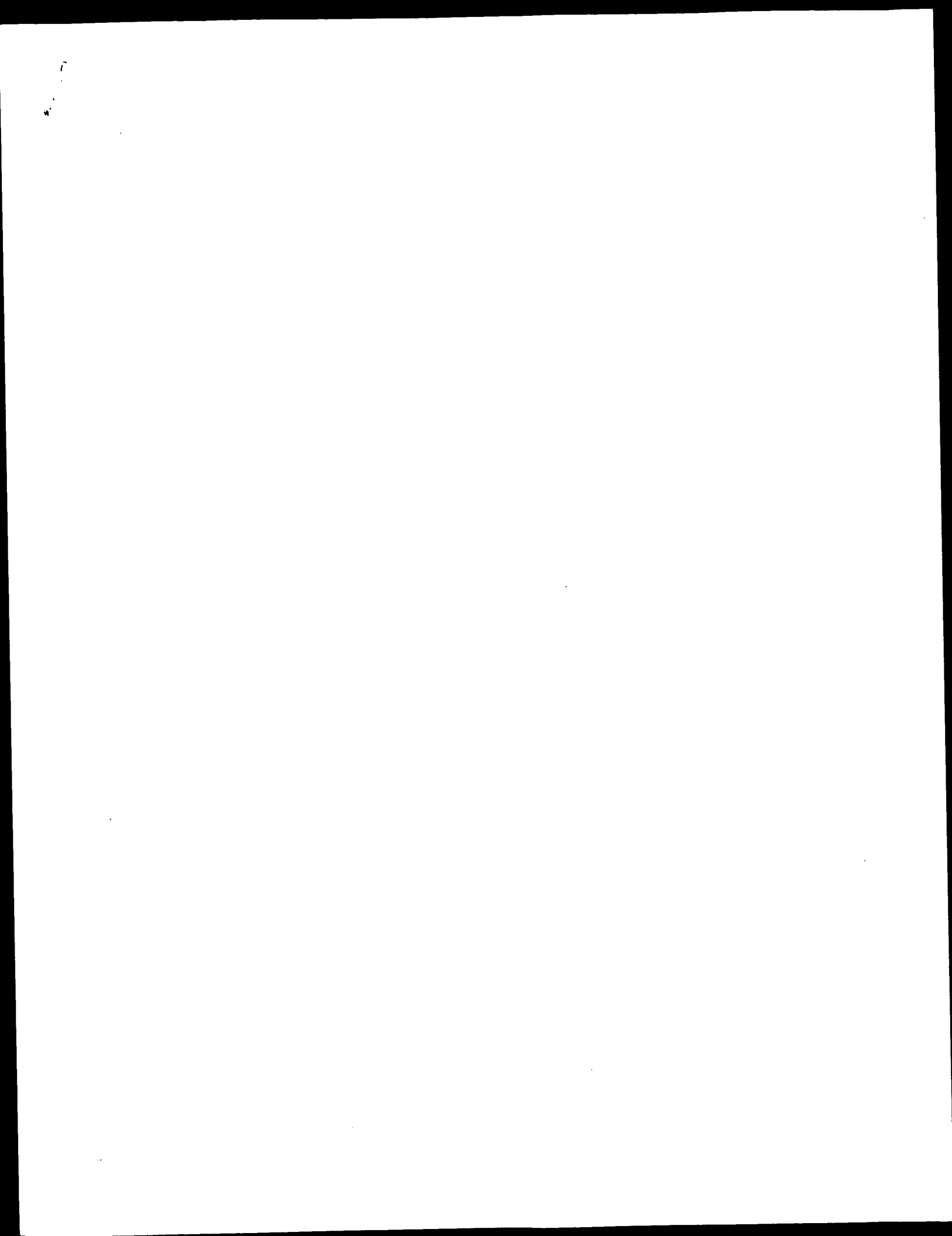
```

; LENGTH: 179
; TYPE: PRT
; ORGANISM: Cucumis melo
US-09-362-642-6

```

Query Match	57.8%	Score 726.5	DB 4	Length 179
Best Local Similarity	72.3%	Pred. No. 1.2e-61		
Matches 128	Conservative 18	Mismatches 28	Indels 3	Gaps 2
QY	10	ATFTGGDASCTMGTCGTGYNLTSTGY-TNTPAALSTVLFENDGAACRSCYEALRCNDGMC	68	
Db	3	ATFTGGDASCTMGTCGTGYNLTSTGYNTPAALSTAFENNGALSCGACFEFLKCANPRMC	62	
QY	69	LPGS--TNTVATATLCPNPALPNDGCMNPPRHFDIAEPAFIQIVYTRAGTIVPYSTYR	126	
Db	63	HPSPCLFITATNFCPCPNPALPNDGCMNLPRTHTFDLAMEFLKIAEYFAGTIGPVSYR	122	
QY	127	VPCVKKGGIRFTTNGHSYFNLVLTYNVAGPGDVOVSATKGSSTFGQPMPSNNMGQNNQ	183	
Db	123	VPCRKGGIRFTTNGHSYFNLVLTYNVAGACDDIVRVSYSKSNMGSMGSMNNMGQNNQ	179	

Search completed: October 11, 2002, 15:02:07
Job time : 8.83985 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 11, 2002, 14:49:55 : Search time 25.0184 Seconds

(without alignments)
1003.367 Million cell updates/sec

Title: US-09-896-301-6
Perfect score: 157
Sequence: 1 DYSSWDSAHATFYGGDASC.....SNNVPAQMGEGTPEGQF 226

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: A_Geneseq_032802.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1246.5	99.2	17	AAR94532	Arabidopsis expans
2	1011.5	80.5	17	AAR94527	Cucurbit expansin
3	960.5	76.4	21	AAG06546	Arabidopsis thalia
4	960.5	76.4	21	AAG51633	Arabidopsis thalia
5	960.5	76.4	21	AAG51647	Arabidopsis thalia
6	960.5	76.4	249	AAG06545	Arabidopsis thalia
7	960.5	76.4	249	AAG51632	Arabidopsis thalia
8	960.5	76.4	249	AAG51646	Arabidopsis thalia
9	960.5	76.4	259	AAG06544	Arabidopsis thalia
10	960.5	76.4	259	AAG51631	Arabidopsis thalia
11	960.5	76.4	280	AAG51645	Arabidopsis thalia

12	959.5	76.3	228	17	AAR94528	Rice expansin. Or
13	948	75.4	253	21	AAG25443	Arabidopsis thalia
14	948	75.4	253	21	AAG46483	Arabidopsis thalia
15	938.5	74.7	251	21	AAG23852	Arabidopsis thalia
16	938.5	74.7	251	21	AAG43343	Arabidopsis thalia
17	938.5	74.7	253	21	AAG23851	Arabidopsis thalia
18	938.5	74.7	253	21	AAG43342	Arabidopsis thalia
19	938.5	74.7	253	22	AAG00412	Tomato seed expans
20	938.5	74.7	281	21	AAG43341	Arabidopsis thalia
21	938.5	74.7	282	21	AAG23850	Arabidopsis thalia
22	932.5	74.2	250	22	AAG00414	Tomato seed expans
23	903.5	71.9	262	21	AAG29931	Arabidopsis thalia
24	903.5	71.9	273	21	AAG29930	Arabidopsis thalia
25	900	71.6	241	21	AAG05453	Arabidopsis thalia
26	900	71.6	249	21	AAG05452	Arabidopsis thalia
27	900	71.6	255	21	AAG05451	Arabidopsis thalia
28	899.5	71.6	257	21	AAG36445	Arabidopsis thalia
29	895	71.2	242	21	AAG36570	Arabidopsis thalia
30	895	71.2	249	21	AAG36569	Arabidopsis thalia
31	895	71.2	255	21	AAG36568	Arabidopsis thalia
32	880.5	70.0	250	21	AAG09622	Arabidopsis thalia
33	880.5	70.0	258	21	AAG09621	Arabidopsis thalia
34	880.5	70.0	280	21	AAG09620	Arabidopsis thalia
35	871	69.3	257	22	AAG00413	Tomato seed expans
36	864.5	68.8	255	21	AAG30325	Arabidopsis thalia
37	864.5	68.8	257	21	AAG30324	Arabidopsis thalia
38	864	68.7	207	21	AAG25444	Arabidopsis thalia
39	864	68.7	207	21	AAG46484	Arabidopsis thalia
40	863.5	68.7	225	17	AAR94531	Arabidopsis thalia
41	816	64.9	221	21	AAG15694	Arabidopsis thalia
42	816	64.9	221	21	AAG51013	Arabidopsis thalia
43	816	64.9	255	21	AAG15693	Arabidopsis thalia
44	816	64.9	255	21	AAG51012	Arabidopsis thalia
45	815.5	64.9	210	21	AAG29932	Arabidopsis thalia

ALIGNMENTS

RESULT 1
ID AAR94532 standard; Protein; 227 AA.
XX
AC AAR94532;
XX
DT 08-JUL-1996 (first entry)
XX
DE Arabidopsis expansin.
XX
KW Expansin: plant cell wall; cellulose; paper recycling; de-linking;
KW polysaccharide.
XX
OS Arabidopsis sp.
XX
PN A09540262-A.
XX
PD 04-APR-1996.
XX
PE 12-MAY-1994; 94AU-0068320.
XX
PR 12-MAY-1995; 95US-0440517.
XX
PR 12-MAY-1993; 93US-0060944.
XX
PA (PENN-) PENN STATE RES FOUND.
XX
PI Cosgrove DJ, McQueen-Mason S;
XX
DR WPI, 1996-201150/21.
XX
PT Expansin proteins which alter the mechanical strength of
XX poly(saccharide(s)) - useful in paper mfr. and recycling
PS Disclosure; Page 32-33; 60pp; English.

XX Expansins are a novel class of proteins that catalyse the extension
 CC of plant cell walls and the weakening of the hydrogen bonds in pure
 CC cellulose. 2 Expansins (AAR94528 and AAR94529) have been identified in
 CC rice and 3 in Arabidopsis (AAR94530-32). A cDNA clone (AAT13320)
 CC coding for cucumber expansin 29 (AAR94527) has been obtained. Expansins
 CC can be used e.g. in the mfr., de-linking and recycling of paper, in
 CC the textile industry, to aid delignification processes, to alter gel
 CC mechanical strength, etc.

XX Sequence 227 AA:

Query Match 99.2%; Score 1246.5; DB 17; Length 227;
 Best Local Similarity 99.6%; Pred. NO. 4.4e-113;
 Matches 226; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 DYSSWQSAHATFYGGDASGTMGTCGYGNLYSTGYTNTALSTVLFNDGACRSCYELR 60
 DB 1 DYSSWQSAHATFYGGDASGTMGTCGYGNLYSTGYTNTALSTVLFNDGACRSCYELR 60
 QY 61 CDNDGQWCLPGSVTVTATNLCPNVALPNDGQWCPNPPHFDMAEPALQIGYRAGIV 120
 DB 61 CDNDGQWCLPGSVTVTATNLCPNVALPNDGQWCPNPPHFDMAEPALQIGYRAGIV 120
 QY 121 PVSYRRVPCYKKGIRFTINGHSYENLVLTNYAGPDVQSVSIKSSSTGQPMRSRNMG 180
 DB 121 PVSYRRVPCYKKGIRFTINGHSYENLVLTNYAGPDVQSVSIKSSSTGQPMRSRNMG 180
 QY 181 NMQSNSTYLDGQSLSFQYAVSDGRTVTSNNVVPAGMFGQTFFG-QGQF 226
 DB 181 NMQSNSTYLDGQSLSFQYAVSDGRTVTSNNVVPAGMFGQTFFG-QGQF 227

RESULT 2
 AAR94527
 ID AAR94527 standard; Protein; 227 AA.

XX AAR94527;

XX 08-JUL-1996 (first entry)

XX Cucurbit expansin-29.

XX Expansin-29; plant cell wall; cellulose; paper recycling; de-linking;
 KM polysaccharide; cucumber.

XX Cucumis sativus var. Burpee Pickler.

XX A09540262-A.

XX 04-APR-1996.

XX 12-MAY-1994; 94AU-0068320.

XX 12-MAY-1995; 95US-0440517.

XX 12-MAY-1993; 93US-0060944.

XX (PENN-) PENN STATE RES FOUND.

XX Cosgrove DJ, McQueen-Mason S;

XX WPI; 1996-201150/21.

XX N-PSDB; AAT13320.

XX Expansin proteins which alter the mechanical strength of
 PT polysaccharide(s) - useful in paper mfr. and recycling

XX Claim 7; Page 30; 60pp; English.

XX Cucurbit expansin-29 (AAR94527) is a member of a novel class of
 CC proteins that catalyse the extension of plant cell walls and the
 CC weakening of the hydrogen bonds in pure cellulose. It can be obtained
 CC by expression of an isolated cDNA clone (see AAT13320) in bacterial or

CC other host cells. Expansin proteins have also been identified in coal
 CC coccidiosis, in Arabidopsis (see AAR94530-32) and in rice (AAR94528-29),
 CC and appear to be broadly distributed throughout the plant kingdom.
 CC Expansins can be used e.g. in the mfr., de-linking and recycling of
 CC paper, in the textile industry, to aid delignification processes, to
 CC alter gel mechanical strength, etc.

XX Sequence 227 AA:

Query Match 80.5%; Score 1011.5; DB 17; Length 227;
 Best Local Similarity 77.1%; Pred. NO. 3.1e-90;
 Matches 175; Conservative 23; Mismatches 28; Indels 1; Gaps 1;

QY 1 DYSSWQSAHATFYGGDASGTMGTCGYGNLYSTGYTNTALSTVLFNDGACRSCYEL 59
 DB 1 DYSSWQSAHATFYGGDASGTMGTCGYGNLYSTGYTNTALSTVLFNDGACRSCYEL 60
 QY 60 RCDNDGQWCLPGSVTVTATNLCPNVALPNDGQWCPNPPHFDMAEPALQIGYRAGI 119
 DB 61 TCTNDPKWCLPGTIRVTATNFCPPNFALPNNNGWCPNPPHFDMAEPALQIGYRAGI 120
 QY 120 VPVSYRRVPCYKKGIRFTINGHSYENLVLTNYAGPDVQSVSIKSSSTGQPMRSRNMG 179
 DB 121 VPVSYRRVPCYKKGIRFTINGHSYENLVLTNYAGPDVQSVSIKSSSTGQPMRSRNMG 180
 QY 180 QNMQSNSTYLDGQSLSFQYAVSDGRTVTSNNVVPAGMFGQTFFG-QGQF 226
 DB 181 QNMQSNSTYLDGQSLSFQYAVSDGRTVTSNNVVPAGMFGQTFFG-QGQF 227

RESULT 3
 AAG06546
 ID AAG06546 standard; Protein; 241 AA.

XX AAG06546;

XX 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 3357.

XX Protein identification; signal transduction pathway; metabolic pathway;
 KM hybridisation assay; genetic mapping; gene expression control; promoter;
 KM termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

XX 05-MAR-1999; 99US-0123180.

XX 09-MAR-1999; 99US-0123548.

XX 23-MAR-1999; 99US-0125788.

XX 25-MAR-1999; 99US-0126264.

XX 29-MAR-1999; 99US-0126785.

XX 01-APR-1999; 99US-0128234.

XX 06-APR-1999; 99US-0128714.

XX 16-APR-1999; 99US-0129845.

XX 19-APR-1999; 99US-0130077.

XX 21-APR-1999; 99US-0130449.

XX 23-APR-1999; 99US-0130891.

XX 28-APR-1999; 99US-0131449.

XX 30-APR-1999; 99US-0132048.

XX 04-MAY-1999; 99US-0132484.

XX 05-MAY-1999; 99US-0132485.

XX 06-MAY-1999; 99US-0132486.

XX 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 24-MAY-1999; 99US-0135353.
PR 25-MAY-1999; 99US-0135629.
PR 27-MAY-1999; 99US-0136021.
PR 28-MAY-1999; 99US-0136392.
PR 01-JUN-1999; 99US-0136782.
PR 03-JUN-1999; 99US-0137222.
PR 04-JUN-1999; 99US-0137528.
PR 07-JUN-1999; 99US-0137502.
PR 08-JUN-1999; 99US-0137724.
PR 10-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139890.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 28-JUN-1999; 99US-0140695.
PR 29-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0140981.
PR 01-JUL-1999; 99US-0141287.
PR 02-JUL-1999; 99US-0141842.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0142977.
PR 14-JUL-1999; 99US-0143542.
PR 15-JUL-1999; 99US-0143624.
PR 16-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144335.
PR 19-JUL-1999; 99US-0144375.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144332.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.

PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147120.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0158294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.

PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 76.4%; Score 960.5; DB 21; Length 241;
Best Local Similarity 75.3%; Pred. No. 3e-85;
Matches 168; Conservative 22; Mismatches 32; Indels 1; Gaps 1;

QY 5 WSAHATEYGGDASGTMGTCGCGNLYSTGT-TWTALSTVLFMDGACACSCYELRCDN 63
I :|||||
DB 18 WINAHATEYGGDASGTMGACGYGNLYSOGYGTSTALSTALFNNGLSGSCSEIRCN 77
I :|||||
QY 64 DQWMLPGSVTATNLCPPNALPNDGCMCPRPPEHMAEPALQIGYRAGIVFS 123
I :|||||
DB 78 DKRWCLPGSIVTATNFCPPNALNNGCMCPLEHFDLAQPVFQRIAGYRAGIVFS 137
I :|||||
QY 124 YRRVPCVKKGGIRFTINGHSYFNLYLVTVNAGPQVSVISIKGSTGMPMSRNNGWQ 183
I :|||||
DB 138 YRRVPCRRGGIRFTINGHSYFNLYLVTVNAGGAGVHSAIIGSRVWQMSRNNGWQ 197
I :|||||
QY 184 SNSYLDGSLSFQVAVSDERTYSNNVPAQMPGOTFEQGF 226
I :|||||
DB 198 SNSYLDGSLSFQVAVSDERTYSNNVPAQMPGOTFEQGF 240
I :|||||

RESULT 4

AAG51633

ID AAG51633 standard; Protein; 241 AA.

XX AAG51633;

XX 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 65551.

XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.

XX Arabidopsis thaliana.

XX EPI033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

XX 05-MAR-1999; 99US-0123180.

XX 09-MAR-1999; 99US-0123548.

XX 23-MAR-1999; 99US-0125788.

XX 25-MAR-1999; 99US-0126264.

XX 29-MAR-1999; 99US-0126785.

XX 01-APR-1999; 99US-0127462.

XX 06-APR-1999; 99US-0128234.

XX 08-APR-1999; 99US-0128714.

XX 16-APR-1999; 99US-0129845.

XX 19-APR-1999; 99US-0130077.

XX 21-APR-1999; 99US-0130449.

XX 23-APR-1999; 99US-0130510.

XX 28-APR-1999; 99US-0130891.

XX 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134270.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140623.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141844.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142053.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.

```

PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145226.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145951.
PR 28-JUL-1999; 99US-0145951.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.

```

```

PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Query Match 76.4%; Score 960.5; DB 21; Length 241;
 Best Local Similarity 75.3%; Pred. No. 3e-85;
 Matches 168; Conservative 22; Mismatches 32; Indels 1; Gaps 1;

```

QY 5 WOSAHATEYGGGDASGTGTCGYNLYSTGY-TTAAALSTVLFNDGAACNSCYELBCDN 63
DB 18 WINAHATEYGGGDASGTGTCGYNLYSTGYSTAAALSTVLFNDGAACNSCYELBCDN 77
QY 64 DQOKLPGSVTYTATNLCPPNVALPNDGCMCNPPRPHDMAEPALDIOGYRACIVPVS 123
DB 78 DQKWCPLPGSVTYTATNLCPPNVALPNDGCMCNPPRPHDMAEPALDIOGYRACIVPVS 133
QY 124 YRRVPCRRRGCIIRFTINGSHSYFNVLVTNVAAGPQVOSYSIKSSSTGMOPMNRNKGMMO 183
DB 138 YRRVPCRRRGCIIRFTINGSHSYFNVLVTNVAAGPQVOSYSIKSSSTGMOPMNRNKGMMO 197
QY 184 SNSYLDGQSLFQVAVSDGRVTSNNVVPAGMFGCTFEBCGF 226
DB 198 SNSYLDGQSLFQVAVSDGRVTSNNVVPAGMFGCTFEBCGF 240

```

RESULT 5
 AAGS1647
 ID AAGS1647 standard; Protein; 241 AA.
 AC AAGS1647;
 XX
 DT 18-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 65570.
 XX
 KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX
 OS Arabidopsis thaliana.
 OS
 PN EP1033405-A2.
 PN
 PD 06-SEP-2000.
 PD
 XX
 XX 25-FEB-2000; 2000EP-0301439.
 PE
 XX 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132477.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134991.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135134.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0138094.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140895.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142927.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 12-AUG-1999; 99US-0148319.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149175.
PR 20-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 07-SEP-1999; 99US-0151930.
PR 10-SEP-1999; 99US-0152363.
PR 13-SEP-1999; 99US-0153070.
PR 15-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154018.
PR 20-SEP-1999; 99US-0154039.
PR 22-SEP-1999; 99US-0154779.
PR 23-SEP-1999; 99US-0155139.
PR 24-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0155659.
PR 29-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158232.
PR 13-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.

Query Match	76.4%;	Score 960.5;	DB 21;	Length 249;
Best Local Similarity	75.3%;	Pred. No. 3,2e-85;		
Matches 168;	Conservative 22;	Mismatches 32;	Indels 1;	Gaps 1
13-OCT-1999;	990S-0159293.	990S-0159294.	990S-0159295.	990S-0159296.
13-OCT-1999;	990S-0159297.	990S-0159298.	990S-0159299.	990S-0159300.
13-OCT-1999;	990S-0159301.	990S-0159302.	990S-0159303.	990S-0159304.
13-OCT-1999;	990S-0159305.	990S-0159306.	990S-0159307.	990S-0159308.
13-OCT-1999;	990S-0159309.	990S-0159310.	990S-0159311.	990S-0159312.
13-OCT-1999;	990S-0159313.	990S-0159314.	990S-0159315.	990S-0159316.
13-OCT-1999;	990S-0159317.	990S-0159318.	990S-0159319.	990S-0159320.
13-OCT-1999;	990S-0159321.	990S-0159322.	990S-0159323.	990S-0159324.
13-OCT-1999;	990S-0159325.	990S-0159326.	990S-0159327.	990S-0159328.
13-OCT-1999;	990S-0159329.	990S-0159330.	990S-0159331.	990S-0159332.
13-OCT-1999;	990S-0159333.	990S-0159334.	990S-0159335.	990S-0159336.
13-OCT-1999;	990S-0159337.	990S-0159338.	990S-0159339.	990S-0159340.
13-OCT-1999;	990S-0159341.	990S-0159342.	990S-0159343.	990S-0159344.
13-OCT-1999;	990S-0159345.	990S-0159346.	990S-0159347.	990S-0159348.
13-OCT-1999;	990S-0159349.	990S-0159350.	990S-0159351.	990S-0159352.
13-OCT-1999;	990S-0159353.	990S-0159354.	990S-0159355.	990S-0159356.
13-OCT-1999;	990S-0159357.	990S-0159358.	990S-0159359.	990S-0159360.
13-OCT-1999;	990S-0159361.	990S-0159362.	990S-0159363.	990S-0159364.
13-OCT-1999;	990S-0159365.	990S-0159366.	990S-0159367.	990S-0159368.
13-OCT-1999;	990S-0159369.	990S-0159370.	990S-0159371.	990S-0159372.
13-OCT-1999;	990S-0159373.	990S-0159374.	990S-0159375.	990S-0159376.
13-OCT-1999;	990S-0159377.	990S-0159378.	990S-0159379.	990S-0159380.
13-OCT-1999;	990S-0159381.	990S-0159382.	990S-0159383.	990S-0159384.
13-OCT-1999;	990S-0159385.	990S-0159386.	990S-0159387.	990S-0159388.
13-OCT-1999;	990S-0159389.	990S-0159390.	990S-0159391.	990S-0159392.
13-OCT-1999;	990S-0159393.	990S-0159394.	990S-0159395.	990S-0159396.
13-OCT-1999;	990S-0159397.	990S-0159398.	990S-0159399.	990S-0159400.
13-OCT-1999;	990S-0159401.	990S-0159402.	990S-0159403.	990S-0159404.
13-OCT-1999;	990S-0159405.	990S-0159406.	990S-0159407.	990S-0159408.
13-OCT-1999;	990S-0159409.	990S-0159410.	990S-0159411.	990S-0159412.
13-OCT-1999;	990S-0159413.	990S-0159414.	990S-0159415.	990S-0159416.
13-OCT-1999;	990S-0159417.	990S-0159418.	990S-0159419.	990S-0159420.
13-OCT-1999;	990S-0159421.	990S-0159422.	990S-0159423.	990S-0159424.
13-OCT-1999;	990S-0159425.	990S-0159426.	990S-0159427.	990S-0159428.
13-OCT-1999;	990S-0159429.	990S-0159430.	990S-0159431.	990S-0159432.
13-OCT-1999;	990S-0159433.	990S-0159434.	990S-0159435.	990S-0159436.
13-OCT-1999;	990S-0159437.	990S-0159438.	990S-0159439.	990S-0159440.
13-OCT-1999;	990S-0159441.	990S-0159442.	990S-0159443.	990S-0159444.
13-OCT-1999;	990S-0159445.	990S-0159446.	990S-0159447.	990S-0159448.
13-OCT-1999;	990S-0159449.	990S-0159450.	990S-0159451.	990S-0159452.
13-OCT-1999;	990S-0159453.	990S-0159454.	990S-0159455.	990S-0159456.
13-OCT-1999;	990S-0159457.	990S-0159458.	990S-0159459.	990S-0159460.
13-OCT-1999;	990S-0159461.	990S-0159462.	990S-0159463.	990S-0159464.
13-OCT-1999;	990S-0159465.	990S-0159466.	990S-0159467.	990S-0159468.
13-OCT-1999;	990S-0159469.	990S-0159470.		

PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125358.
PR 25-MAR-1999; 99US-0125788.
PR 29-MAR-1999; 99US-0126264.
PR 01-APR-1999; 99US-0126785.
PR 06-APR-1999; 99US-0127462.
PR 08-APR-1999; 99US-0128234.
PR 16-APR-1999; 99US-0128714.
PR 19-APR-1999; 99US-0129845.
PR 21-APR-1999; 99US-0130077.
PR 23-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132556.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140891.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142305.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142820.
PR 13-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148584.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151348.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.


```

PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151530.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.

```

```

PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Query Match 76.4%; Score 960.5; DB 21; Length 249;
 Best Local Similarity 75.3%; Pred. No. 3, 2e-85;
 Matches 168; Conservative 22; Mismatches 32; Indels 1; Gaps 1;

```

QY 5 WSAHATFYGGDASGMCCTGCGNLSXGY--TNTAALSTVFENNGACRSCELRKDN 63
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 26 WINAHATFYGGDASGTMGACGYGNLSGYGTSTPAALSTALENNGLSGSCFETRCEN 85
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 64 DQWCLPGSVYVATNLCPNPVYALPNDGGMCPNPPHFDMAEPATLOIGVYRAGIVPV 123
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 86 DGKMCPLPGSIVYATNCPNPNMNLANNNGCMCPNPLEHFDIAQPVFORIAOTYRAGIVPV 145
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 124 YRRVPCYKKGKIRFTINGHSYFNLVTVNAGFGDVQSVSIKSSYTGWQPMNRNMGQNN 183
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 146 YRRVPCRRRCGIRFTINGHSYFNLVTVNAGAGDVHSAIKSRVWQAMSRNMGQNN 205
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 184 SNSLYDGOSSLFOVAVSDGRTVSNNVVPAGMFGOTFEFGCOR 226
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 206 SNSLYNGOALSFVYTTSDGRTVVSFNAAPAGMSYGTOTFAGGCP 248
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

RESULT 9
 AAG06544
 ID AAG06544 standard; Protein; 259 AA.
 XX
 AC AAG06544;
 XX
 DT 17-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 3355.
 XX
 KW protein identification; signal transduction pathway; metabolic pathway;
 hybridisation assay; genetic mapping; gene expression control; promoter;
 termination sequence.

XX Arabidopsis thaliana.
OS
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0122180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 23-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135529.
PR 24-MAY-1999; 99US-0136021.
PR 25-MAY-1999; 99US-0136592.
PR 27-MAY-1999; 99US-0136782.
PR 28-MAY-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 18-JUN-1999; 99US-0139817.
PR 21-JUN-1999; 99US-0139839.
PR 22-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.

PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0148426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.

PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 76.4%; Score 960.5; DB 21; length 259;
Best Local Similarity 75.3%; Pred. No. 3.3e-85;
Matches 168; Conservative 22; Mismatches 32; Indels 1; Gaps 1;

QY 5 WQSHATFFYGGGASGTMGTCGYNLYSTGY-TNTAALSTVLEPNDGACRSCYEELRCDN 63
DB 36 WINHAATFYGGGASGTMGACGYNLYSOGYGTSTAAALSTALFNNGLSCGSCFEIRCN 95
QY 64 DGQWCLPGSVTVATNATNCPPNYALPNDGCMCPRPHPMAEPALQIOTIVYRAGITVPS 123
DB 96 DGKWCCLPGSLVTVATNATNCPPNNALANNNGCMCPLEHFDLADPVORIKQIRAGITVPS 155
QY 124 YRRVPCVKKKGIRFTINGHSYFNLVLTNNVAGGQVQSVISIKSSGTGQWPMRNNQNMQ 183
DB 156 YRRVPCRRRGIRFTINGHSYFNLVLTNNVAGGADVHSAIKSRTVWQMSRNNQNMQ 215
QY 184 SNSYLDGOSLFOVAVSDGRTVTSNNVPPAGMOPGQTFEEGQF 226
DB 216 SNSYLDGOSLFOVAVSDGRTVTSNNVPPAGMOPGQTFEEGQF 258

RESULT 10
AAG51631
ID AAG51631 standard; protein; 259 AA.
XX AAG51631:
AC
XX
DT 18-OCT-2000 (first entry)
XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 65549.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 03-JUN-1999; 99US-0137222.
PR 04-JUN-1999; 99US-0137528.
PR 07-JUN-1999; 99US-0137502.
PR 08-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140999.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144684.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145513.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 18-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149729.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.

PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0156659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 13-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161049.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 26-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 76.4%; Score 960.5; DB 21; Length 259;
 Best Local Similarity 75.3%; Pred. No. 3.3e-85;
 Matches 168; Conservative 22; Mismatches 32; Indels 1; Gaps 1;

QY 5 MOSAHATFYGGDASCTMGCTGCGNYLSTGY-TTAAISTVLFPNDGAACRSCELRCDN 63
 DB 36 WTAHATFTYGGDASCTMGACGNYLSTGYTAAISTVLFPNDGAACRSCELRCDN 95
 QY 64 DQWCLPGSVTYATNLCPPNYALPNDGWCNPPRPHEDMAEPFLDIGYRAGIATVVS 123
 DB 96 DQWCLPGSVTYATNLCPPNYALPNDGWCNPPRPHEDMAEPFLDIGYRAGIATVVS 155
 QY 124 YRRVPCRRRGIRFTINGHSYFNLYLVYTNVAGPGVQSVSTIGSGTGMOPSRNNGOMNO 183
 DB 156 YRRVPCRRRGIRFTINGHSYFNLYLVYTNVAGPGVQSVSTIGSGTGMOPSRNNGOMNO 215
 QY 184 SNSYLDGSLSPVAVSDGRVTSSNNVVPAGMOFCQTEFGCOF 226
 DB 216 SNSYLDGSLSPVAVSDGRVTSSNNVVPAGMOFCQTEFGCOF 258

RESULT 11
 AAG51645
 ID AAG51645 standard; Protein; 280 AA.

XX AC AAG51645;
XX 18-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 65568.
XX
XX Protein identification: signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EPI033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126284.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130891.
XX 30-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 04-MAY-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 07-MAY-1999; 99US-0132487.
XX 07-MAY-1999; 99US-0132863.
XX 11-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 17-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139492.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
XX 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.
XX 18-JUN-1999; 99US-0139460.
XX 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139849.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 12-AUG-1999; 99US-0148319.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149724.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.

PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159285.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.

	Query Match	Similarity	75.3%	Score	960.5	DB	Length	280
	Best Local	Similarity	75.3%	Pred. No.	3.7e-85			
	Matches	168	Conservative	22	Mismatches	32	Indels	1
QY	5	WOSAHATEYGGGADSGTMTGTCGTYGMLYSTGY-TNTALSTVLEFNDGAACRCYELRCDN	63					
DB	57	WINAHATEYGGGADSGTMTGACGTYGMLYSTGYSTALSTFLFNGTLCSCGSCFEELRCN	116					
QY	64	DGOMKLPGSVYVYATNLCPENALPDPDGGCMKNPPPHFDMAEPALQIGYVRAGTIVPS	123					
DB	117	DGKMCLPGSIVYVYATNLCPENALPDPDGGCMKNPPPHFDMAEPALQIGYVRAGTIVPS	176					
QY	124	YRRVPCVKKGGTRETNGHSYFNLVLYTVNVAAGPDGVSYISGSSGTGWDPMSNMGNMNO	183					
DB	177	YRRVPCRRRGRIRETNGHSYFNLVLYTVNVAAGPDGVSAIKGSRVTWQAMSRMGNMNO	236					
QY	184	SNSYLDGOSLFOQAVASDGRVTSNNVYVAGMQFGTFFGGGF	226					
DB	237	SNSYTLNGQALSERKVTTSDDGRIVTSFPAADAGSYGCTFFGGGF	279					

Sequence	228 AA:
XX	RESULT 12
AC	AAR94528
ID	AAR94528 standard; Protein; 228 AA.
XX	AAR94528;
XX	
DT	08-JUL-1996 (first entry)
XX	
DE	Rice expansin.
XX	
KM	Expansin; plant cell wall; cellulose; paper recycling; de-linking;
KM	polysaccharide; rice.
XX	
OS	Oryza sativa.
XX	
FH	Key Location/Qualifiers
FT	Misc-difference 211
FT	/note= "unidentified amino acid"
XX	
PN	AU9540262-A.
XX	
PD	04-APR-1996.
XX	
PE	12-MAY-1994; 94AU-0068320.
XX	
PR	12-MAY-1995; 95US-0440517.
PR	12-MAY-1993; 93US-0060944.
XX	
PA	(PENN-) PENN STATE RES FOUND.
XX	
PI	Cosgrove DJ, McQueen-Mason S;
XX	
DR	WPI; 1996-201150/21.
XX	
PT	Expansin proteins which alter the mechanical strength of
XX	poly(saccharide(s) - useful in paper mfr. and recycling
XX	
PS	Disclosure; Page 30-31; 60pp; English.
XX	
CC	Expansins are a novel class of proteins that catalyse the extension
CC	of plant cell walls and the weakening of the hydrogen bonds in pure
CC	cellulose. 2 Expansins (AAR94528 and AAR94529) have been identified in
CC	rice and 3 in Arabidopsis (AAR94530-32). A cDNA clone (AA113320)
CC	coding for cucumber expansin 29 (AAR94527) has been obtd. Expansins
CC	can be used e.g. in the mfr., de-linking and recycling of paper, in
CC	the textile industry, to aid delignification processes, to alter gel
XX	mechanical strength, etc.
XX	

[illegible]

RESULT 13
AAG25443
ID AAG25443 standard; Protein; 253 AA.
XX
AC AAG25443;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 29510.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140635.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145091.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145224.
PR 23-JUL-1999; 99US-0145228.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.

```
PR 20-AUG-1999; 9905-0149929.
PR 23-AUG-1999; 9905-0149902.
PR 23-AUG-1999; 9905-0149930.
PR 25-AUG-1999; 9905-0150566.
PR 26-AUG-1999; 9905-0150884.
PR 27-AUG-1999; 9905-0151085.
PR 27-AUG-1999; 9905-0151066.
PR 27-AUG-1999; 9905-0151080.
PR 30-AUG-1999; 9905-0151303.
PR 31-AUG-1999; 9905-0151438.
PR 01-SEP-1999; 9905-0151930.
PR 07-SEP-1999; 9905-0152363.
PR 10-SEP-1999; 9905-0153070.
PR 13-SEP-1999; 9905-0153758.
PR 15-SEP-1999; 9905-0154018.
PR 16-SEP-1999; 9905-0154039.
PR 20-SEP-1999; 9905-0154779.
PR 22-SEP-1999; 9905-0155139.
PR 23-SEP-1999; 9905-0155486.
PR 24-SEP-1999; 9905-0155659.
PR 28-SEP-1999; 9905-0156458.
PR 29-SEP-1999; 9905-0156596.
PR 04-OCT-1999; 9905-0157117.
PR 05-OCT-1999; 9905-0157753.
PR 06-OCT-1999; 9905-0157865.
PR 07-OCT-1999; 9905-0158029.
PR 08-OCT-1999; 9905-0158232.
PR 12-OCT-1999; 9905-0158369.
PR 13-OCT-1999; 9905-0159293.
PR 13-OCT-1999; 9905-0159284.
PR 13-OCT-1999; 9905-0159295.
PR 14-OCT-1999; 9905-0159329.
PR 14-OCT-1999; 9905-0159330.
PR 14-OCT-1999; 9905-0159331.
PR 14-OCT-1999; 9905-0159637.
PR 14-OCT-1999; 9905-0159638.
PR 18-OCT-1999; 9905-0159584.
PR 21-OCT-1999; 9905-0160741.
PR 21-OCT-1999; 9905-0160767.
PR 21-OCT-1999; 9905-0160768.
PR 21-OCT-1999; 9905-0160770.
PR 21-OCT-1999; 9905-0160814.
PR 21-OCT-1999; 9905-0160815.
PR 22-OCT-1999; 9905-0160980.
PR 22-OCT-1999; 9905-0160981.
PR 22-OCT-1999; 9905-0160989.
PR 25-OCT-1999; 9905-0161404.
PR 25-OCT-1999; 9905-0161405.
PR 25-OCT-1999; 9905-0161406.
PR 26-OCT-1999; 9905-0161359.
PR 26-OCT-1999; 9905-0161360.
PR 26-OCT-1999; 9905-0161361.
PR 28-OCT-1999; 9905-0161920.
PR 28-OCT-1999; 9905-0161992.
PR 28-OCT-1999; 9905-0161993.
PR 29-OCT-1999; 9905-0162142.
```

Query Match 75.4%; Score 948; DB 21; Length 253;

Best Local Similarity 73.2%; Pred. No. 5, 3e-84; Mismatches 36; Indels 2; Gaps 2;

Matches 167; Conservative 23; Mismatches 36; Indels 2; Gaps 2;

```
QY 1 DYSSWQSAHATFFYGGDASGTGTCGYGNLYSTGY-TNTAALSTVLEFNDGAACRSCYEL 59
DB 26 DDGGMQGHATFFYGGDASGTGTCGYGNLYSTGY-TNTAALSTVLEFNDGAACRSCYEL 85
QY 60 RCDNDGOMCLPGSVTVATNLCPNPYALPNDGGMCPRPPEHDMAPALQIGYRAGI 119
DB 86 KCNDPDMWCLPGSVTVATNLCPNPYALPNDGGMCPRPPEHDMAPALQIGYRAGI 145
QY 120 VPVSYRRVPCVKKGIFFTNGHSYFNLVLTNNVAGPGDVOSYIKSSST-GMOPMSRN 178
DB 146 VPVSYRRVPCVKKGIFFTNGHSYFNLVLTNNVAGPGDVOSYIKSSST-GMOPMSRN 205
```

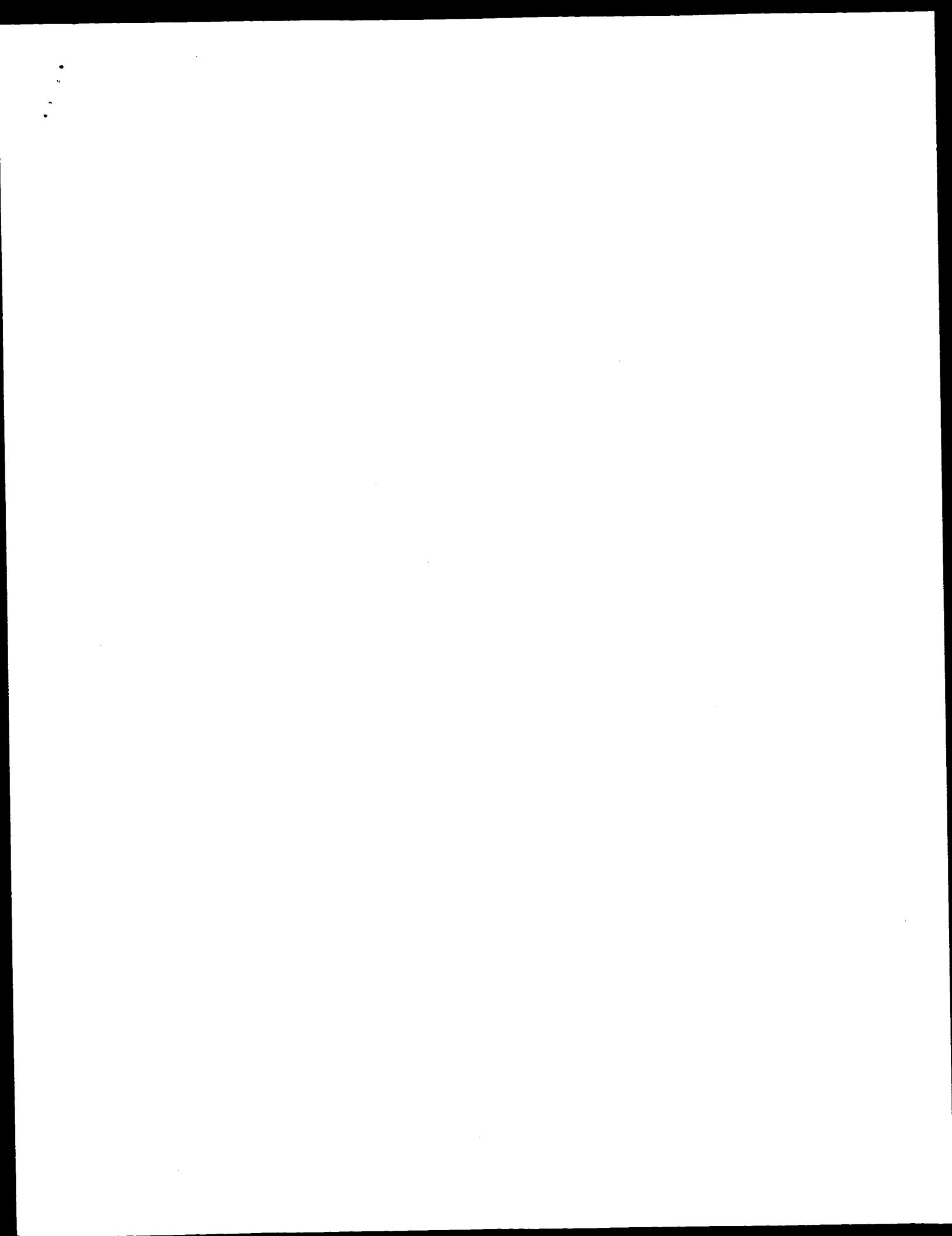
```
QY 179 GQMWQNSYLDGSLSEFOVAVSDGRTVTSNNVYPAGMGFGQTEEGQF 226
DB 206 GQMWQNSYMWNDQSLSEFOVAVSDGRTVTSNNVYPAGMGFGQTEEGQF 253
RESULT 14
AAG46483
AAG46483 standard; Protein; 253 AA.
XX
AC AAG46483;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 58485.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 9905-0121825.
PR 05-MAR-1999; 9905-0123180.
PR 09-MAR-1999; 9905-0123548.
PR 23-MAR-1999; 9905-0125788.
PR 25-MAR-1999; 9905-0126264.
PR 29-MAR-1999; 9905-0126785.
PR 01-APR-1999; 9905-0127462.
PR 06-APR-1999; 9905-0128234.
PR 08-APR-1999; 9905-0128714.
PR 16-APR-1999; 9905-0128845.
PR 19-APR-1999; 9905-0130077.
PR 21-APR-1999; 9905-0130449.
PR 23-APR-1999; 9905-0130510.
PR 23-APR-1999; 9905-0130891.
PR 28-APR-1999; 9905-0131449.
PR 30-APR-1999; 9905-0132048.
PR 30-APR-1999; 9905-0132407.
PR 04-MAY-1999; 9905-0132484.
PR 05-MAY-1999; 9905-0132485.
PR 06-MAY-1999; 9905-0132486.
PR 06-MAY-1999; 9905-0132487.
PR 07-MAY-1999; 9905-0132863.
PR 11-MAY-1999; 9905-0134256.
PR 14-MAY-1999; 9905-0134218.
PR 14-MAY-1999; 9905-0134219.
PR 14-MAY-1999; 9905-0134221.
PR 14-MAY-1999; 9905-0134370.
PR 18-MAY-1999; 9905-0134768.
PR 19-MAY-1999; 9905-0134941.
PR 20-MAY-1999; 9905-0135124.
PR 21-MAY-1999; 9905-0135353.
PR 24-MAY-1999; 9905-0135629.
PR 25-MAY-1999; 9905-0136021.
PR 27-MAY-1999; 9905-0136392.
PR 01-JUN-1999; 9905-0136782.
PR 03-JUN-1999; 9905-0137528.
PR 04-JUN-1999; 9905-0137502.
PR 07-JUN-1999; 9905-0137724.
PR 08-JUN-1999; 9905-0138094.
PR 10-JUN-1999; 9905-0138540.
PR 10-JUN-1999; 9905-0138847.
PR 14-JUN-1999; 9905-0139119.
PR 16-JUN-1999; 9905-0139452.
PR 17-JUN-1999; 9905-0139453.
PR 17-JUN-1999; 9905-0139492.
```


PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142927.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.

PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149729.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157855.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 23-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 75.4%; Score 948; DB 21; Length 253;
Best Local Similarity 73.2%; Pred. No. 5,36-84;
Matches 167; Conservative 23; Mismatches 36; Indels 2; Gaps 2;

Oy 1 DYSSWQSAHATFYGGGDSGTMGTGCGYGNLYSTGY-TNFAALSTVLJPNDGAACRSCYEL 59
Db 26 DDGNGQGHATFYGGEDSDSGTMGACGYNLYXGCGYGNLTALSTALFPNNGLTGACCEM 85
Oy 60 RCDNDGQWCLGPGSVTVATNLCPNYALPNDGQWGNPPRPFMADEPAFLQIGVYRAGI 119
Db 86 KCNDPFRWCLGSTITVTATNCPNPFGISNDNGWGNPPLQHPDLAEPALQIAOYRAGI 145



GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 11, 2002, 14:51:00 ; Search time 19.6007 Seconds
(without alignments)
2003.488 Million cell updates/sec

Title: US-09-896-301-7
Perfect score: 1275
Sequence: 1 DYGWQSGHATFGGAGDASG.....AYNLVPSNMWEGQTEGPOF 227

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SP-archaea:*
2: SP-bacteria:*
3: SP-fungi:*
4: SP-human:*
5: SP-invertebrate:*
6: SP-mammal:*
7: SP-mhc:*
8: SP-organella:*
9: SP-phage:*
10: SP-plant:*
11: SP-rodent:*
12: SP-virus:*
13: SP-vertebrate:*
14: SP-unclassified:*
15: SP-virus:*
16: SP-bacteriophage:*
17: SP-archaeop:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1265	99.2	250	10	Q39625 cucumis sat
2	1143	89.6	254	10	Q82093 prunus arne
3	1138	89.3	245	10	Q91LB2 zinnia eleg
4	1134	88.9	254	10	Q9FNU3 prunus aviu
5	1134	88.9	260	10	Q9FNU0 prunus arlet
6	1105.5	86.7	253	10	Q9SBR1 fragaria an
7	1103.5	86.5	252	10	Q93XP2 prunus cera
8	1101	86.4	252	10	Q9FS30 prunus pers
9	1095	85.9	232	10	Q81133 prunus arne
10	1092	85.6	253	10	Q93493 prunus taeda
11	1087	85.3	232	10	Q9SWY1 prunus taeda
12	1087	85.3	232	10	P93492 prunus taeda
13	1085	85.1	232	10	P93495 prunus taeda
14	1084	85.0	247	10	Q82625 lycopersico
15	1081.5	84.8	253	10	Q9SW44 rumex palus
16	1077	84.5	232	10	P93494 prunus taeda

17	1072.5	84.1	253	10	Q22874 arabidopsis
18	1071	84.0	249	10	Q92P35 nicotiana t
19	1056	82.8	251	10	Q94671 oryza sativ
20	1042	81.7	246	10	Q94600 prunus cera
21	1042	81.7	249	10	Q93XP1 prunus cera
22	1041	81.6	246	10	P93442 oryza sativ
23	1041	81.6	251	10	Q40636 oryza sativ
24	1041	81.6	258	10	Q49194 gossypium h
25	1037	81.3	253	10	Q94K72 zea mays (m
26	1032	80.9	252	10	Q9FNU2 prunus avu
27	1031	80.9	249	10	Q9LDR9 arabidopsis
28	1027	80.5	237	10	Q38863 arabidopsis
29	1027	80.5	250	10	Q9CS54 arabidopsis
30	1026	80.5	247	10	Q9M517 trichysaria
31	1022.5	80.2	255	10	Q38866 arabidopsis
32	1019	79.9	248	10	Q80622 arabidopsis
33	1017	79.8	249	10	Q92P36 arabidopsis
34	1016.5	79.7	255	10	Q9FLC5 arabidopsis
35	1014.5	79.6	249	10	Q9FMA0 arabidopsis
36	997.5	78.2	242	10	Q9LBI1 zinnia eleg
37	997	78.2	250	10	Q9FVQ9 lycopersico
38	993	77.9	220	10	Q9AYR1 eustoma gra
39	985	77.3	257	10	Q48818 arabidopsis
40	967	75.8	239	10	Q92P31 lycopersico
41	960	75.3	255	10	Q9FMA0 arabidopsis
42	960	75.3	260	10	Q9M2S9 arabidopsis
43	958	75.1	257	10	Q9SD24 marsilea qu
44	954	74.8	262	10	Q80932 arabidopsis
45	953.5	74.8	252	10	Q9FY30 festuca pra

ALIGNMENTS

RESULT 1
Q39625 PRELIMINARY; PRT; 250 AA.
AC Q39625;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE EXPANSIN S1.
GN CS-EXPI.
OS Cucumis sativus (Cucumber).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Cucurbitales; Cucurbitaceae; Cucumns.
OX NCBI_TaxID=3659;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BURPEE PICKLER;
RX MEDLINE=96016146; PubMed=7568110;
RA Shoberhan T.Y., Shi J., Duracko D.M., Gultinan M.J.,
McQueen-Mason S.J., Shieh M., Cosgrove D.J.;
RT "Molecular cloning and sequence analysis of expansin--a highly
RT conserved, multigene family of proteins that mediate cell wall
RT extension in plants.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:9245-9249(1995).
DR EMBL; U30382; AAB37746.1; -;
DR InterPro; IPR000882; Pollen_allergen.
DR Pfam; PF01357; Pollen_allergen; 1.
DR PRINTS; PR01225; EXPANSIN_FAMLY.
DR PRODOM; PD002179; Pollen_allergen; 1.
FT CHAIN 24
FT SEQUENCE 250 AA; 27215 MW; 60651BC47E4186DA CRC64;

Query Match 99.2%; Score 1265; DB 10; Length 250;
Best Local Similarity 99.1%; Pred. No. 1.5e-108;
Matches 225; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYGWQSGHATFGGAGDASGTCGACGCGNLVSGCGTNTVALSTALFNNGLSGCACFEM 60
DB 24 DYGWQSGHATFGGAGDASGTCGACGCGNLVSGCGTNTVALSTALFNNGLSGCACFEM 83

QY	61	TCTNDPKKCLEGGTIRVATATNCPCPPAFALPNDDGGCNCPLQHFDMAEPAFLQIAOYRAGI	120
DB	84	TCTNDPKKCLEGGTIRVATATNCPCPPAFALPNDDGGCNCPLQHFDMAEPAFLQIAOYRAGI	143
QY	121	VPVSRFRVPCMKKGGVRFRTINGHSYFNVLITTNVGGADVHSVSIKGSRTGQMSRRNG	180
DB	144	VPVSRFRVPCMKKGGVRFRTINGHSYFNVLITTNVGGADVHSVSIKGSRTGQMSRRNG	203
QY	181	QNMOSNNYTLNGGSLFQVTLSDGRTLTAYNLVPSSMWGQGTTEGPOF	227
DB	204	QNMOSNNYTLNGGSLFQVTLSDGRTLTAYNLVPSSMWGQGTTEGPOF	250
RESULT 2			
ID	082093	PRELIMINARY:	PRT: 254 AA.
AC	082093:		
DT	01-NOV-1998 (TREMBLrel. 08, Created)		
DT	01-NOV-1998 (TREMBLrel. 08, Last sequence update)		
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)		
DE	EXPANSIN.		
GN	PA-EXPI.		
OS	Prunus armeniaca (Apricot).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;		
OC	eurosidia I; Rosales; Rosaceae; Amygdaloideae; Prunus.		
OX	NCBI_TaxID=36596;		
RN	[1]		
RP	SEQUENCE FROM N. A.		
RC	STRAIN=BERGERON; TISSUE=MESOCARP PLUS EXOCARP;		
RA	Mbeugle-A-Mbeugle D., Gomez R.-M., Fils-Lycoun B.;		
RT	"Molecular cloning and nucleotide sequence of expansin 1 (PA-Exp1)		
RT	from apricot fruit ";		
RL	Submitted (DEC-1997) to the EMBL/Genbank/DBJ databases.		
DR	EMBL; U93167; AAC3529.1; -		
DR	InterPro; IPR000882; Pollen_allergen.		
DR	Pfam; PF01357; Pollen_allergen; 1.		
DR	PRINTS; PR01225; EXPANSINFAMILY.		
DR	Prodom; PD002179; Pollen_allergen; 1.		
SO	SEQUENCE 254 AA: 27264 MW: 88068D75932FD0E1 CRC64:		
Query Match 89.6%; Score 1143; DB 10; Length 254;			
Best Local Similarity 87.2%; Pred. No. 2.5e-97;			
Matches 198; Conservative 13; Mismatches 16; Indels 0; Gaps 0.			
QY	1	DVGGMOSGATRYGGSDASGTIMGAGCGYNLVSOGGTNTVTLSTALFFNNGLSGCGCFEM	60
DB	28	DVGGMGATATRYGGSDASGTIMGAGCGYNLVSOGGTNTVTLSTALFFNNGLSGCGCFEM	87
QY	61	TCTNDPKKCLEGGTIRVATATNCPCPPAFALPNDDGGCNCPLQHFDMAEPAFLQIAOYRAGI	120
DB	88	TCTNDPKKCLEGGTIRVATATNCPCPPAFALPNDDGGCNCPLQHFDMAEPAFLQIAOYRAGI	147
QY	121	VPVSRFRVPCMKKGGVRFRTINGHSYFNVLITTNVGGAGVHSVSIKGSRTGQMSRRNG	180
DB	148	VPVSRFRVPCMKKGGVRFRTINGHSYFNVLITTNVGGAGVHSVSIKGSRTGQMSRRNG	207
QY	181	QNMOSNNYTLNGGSLFQVTLSDGRTLTAYNLVPSSMWGQGTTEGPOF	227
DB	208	QNMOSNNYTLNGGSLFQVTLSDGRTLTAYNLVPSSMWGQGTTEGPOF	254
RESULT 3			
ID	091LB2	PRELIMINARY:	PRT: 245 AA.
AC	091LB2:		
DT	01-OCT-2000 (TREMBLrel. 15, Created)		
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)		
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)		
DE	EXPANSIN 2.		
OS	Zinnia elegans.		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		

Query Match	89.3%;	Score 1138;	DB 10;	Length 245;
Best Local Similarity	86.7%;	Pred. No. 7e-97;		
Matches 196;	Conservative 19;	Mismatches 11;	Indels 0;	Gaps 0;
QY 2 YGGWMSGHATFEYGGGDASGTMGAGCGYGLYSGQGTNTVALSTALFNNGLSCAGFEWT 61				
DB 20 YGGWMSGHATFEYGGGDASGTMGAGCGYGLYSGQGTNTVALSTALFNNGLSCAGFEWT 79				
QY 62 CTNDRKWCILPCTIRYATNFCPPNFALPNDGSGWCNPLQHFMDAPFLQIAQYRACIV 121				
DB 80 CNDDEPRKWCILPCTIRYATNFCPPNFALPNDGSGWCNPLQHFMDAPFLQIAQYRACIV 139				
QY 122 PVSFRRVPCMKRGCVRTFNHGSYNTNVLITVWGAGDVHVSISGRFGWMSNRNMQ 181				
DB 140 PVSFRRVPCMKRGCVRTFNHGSYNTNVLITVWGAGDVHVSISGRFGWMSNRNMQ 199				
QY 182 NMQSNMYLNGQSLSFQVLTSDGRITLAVNLVPSNMQFGQTYEGPOF 227				
DB 200 NMQSNMYLNGQSLSFQVLTSDGRITLAVNLVPSNMQFGQTYEGPOF 245				
RESULT 4				
ID Q9FUM3 PRELIMINARY; PRT; 254 AA.				
AC Q9FUM3;				
DT 01-MAR-2001 (TREMBLrel. 16, Created)				
DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)				
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)				
DE EXPANSIN 1.				
GN Exptl OR EXP2.				
OS Prunus avium (Cherry), and				
OS Prunus cerasus.				
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;				
OC eucosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.				
OX NCBI_TaxID=42229, 140311;				
RN [1]				
RP SEQUENCE FROM N.A.				
RP SPECIES=P. avium;				
RC SPECIES=P. avium;				
RA Wu Z., Wiersma P. A.;				
RT "Differential Expression of Expansin Genes Isolated from Sweet Cherry				
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.				
RN [2]				
RP SEQUENCE FROM N.A.				
RP SPECIES=P. cerasus; TISSUE=RIPENING FRUIT;				
RC SPECIES=P. cerasus; TISSUE=RIPENING FRUIT;				
RA Yoo S.-D., Gao Z., Cantini C., Loeschner W., van Nocker S.;				
RT "Coordinated expression of genes encoding xpannsins and other cell				
RL wall-modifying enzymes is associated with pectin-related changes in				
RT the cell wall during ripening of cherry (P. cerasus) fruit.";				
RL Submitted (Feb-2001) to the EMBL/GenBank/DBJ databases.				
DR EMBL: AF297521; AAG13982.1; "				
DR EMBL: AF350937; AAK4846.1; "				
DR InterPro: IPR000882; Pollen_allergen.				
DR Pfam: PF01357; Pollen_allergen. 1.				
DR PRINTS: PRO1225; EXPANSINFAMILY.				

DR ProDom; PD002179; Pollen_allergen; 1.
SQ SEQUENCE 254 AA; 27278 MW; 953A7EB2491FD0E1 CRC64

Query Match	88.9%	Score 1134;	DB 10;	Length 254;
Best Local Similarity	86.8%	Pred. No. 1.7e-96;		
Matches 197; Conservative	12;	Mismatches 18;	Indels 0;	Gaps 0;

QY	1	DYGMQSGHAFYFGGSDASGTMGACGAGCNLYSGGYGTNTWALSTALFNNGLSGCAAFEM	60
		: : : : : : : : : : :	
Db	28	DYGMGEGHNAHFYFGGSDASGTMGACGAGCNLYSGGYGTNTWALSTALFNNGLSGSGCYEM	87
QY	61	TCYNDPKWCKLGRITRYATATNCPENFALPNDGGMGNPRLQHEHMAEPALQIAQYRAGI	120
		: : : : : : : : : : :	
Db	88	RCNNDPRMCRGPGSIIVATNCCPENFAGSNDNGCMGNPRLQHFPLAEPAFLQIAQYRAGI	147
QY	121	VPVSFRFRPCKMKKGGVRRITNGHSYFNLYLTNNGGAGDVHSYISKSRFGWQMSMRNMG	180
		: : : : : : : : : : :	
Db	148	VPVIFRRPCKMKKGGIRFTNGHSYFNLYLTNNGGAGDVHSYISKSRFGWQMSMRNMG	207
QY	181	QNMGSNNYLNNGGLSFQYVTLSDGRITLAVNLPVPSMGFGQYEEGPOF	227
		: : : : : : : : : : :	
Db	208	QNMGSNTYLNNGGLSFQYVTLSDGRITVINYAVAPNMDFGQTFESGQF	254

RESULT	5
09FNT0	
ID	09FNT0
AC	PRELIMINARY;
AC	09FNT0;
DT	01-MAR-2001 (TREMBlrel. 16, Created)
DT	01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT	01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE	EXPANSIN.
OS	Cicer arietinum (Chickpea) (Garbanzo).
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC	Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots; Rosidae
OC	eurosid 1; Fabales; Fabaceae; Papilionoideae; Ciceraceae; Cicer.
ON	NCHI_TaxID=3827;
ON	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=CV. CASTELLANA; TISSUE=ETIOLATED EPICOTYLS;
RC	Doplico B., Sanchez M.A., Labrador E.;
RT	"An second expansin is expressed in chickpea epicotyls.";
RL	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR	EMBL: A0291817; GAC19184.1; -
DR	InterPro: IPR000882; pollen_allergen.
DR	InterPro: IPR000504; RRM.
DR	PRINTS: PR01225; EXPANSIN.FAMILY.
DR	Prodom: PD0003179; pollen_allergen; 1.
DR	PROSITE: PS00030; RRM_RRP_1; UNKNOWN_1
SO	SEQUENCE 260 AA; 28266 MW; 32A7036BED2832E9 CRC64;

Query Match	88.9%	Score 1134;	DB 10;	Length 260;
Best Local Similarity	86.3%	Pred. NO. 1.8e-96;		
Matches 196; Conservative	18;	Mismatches 13;	Indels 0;	Gaps 0;

QY	1	DYGMOSCHHAFYVGGGASSTGGAGXGNLSYSGYCTNTVAALSTALFNNGLSGCAFEEM	60
Db	34	DYGMCHGAHAFYVGGGASSTGGAGCGTGNLSYSGYCTNTVAALSTALFNNGLSGSCYEM	93
QY	61	TCTNDPKMKCGTGTARVATATNCPNPFAALPNDGDMCNPRLQHEMDAPAELOIAOYFAGI	120
Db	94	RCNDPRHMKCGSLIVATATNCPNPBSLANNNGMCNPRQLQHEMDAPAELOIAEYFAGI	153
QY	121	VPVSEFRVPCMKKCGVRFITNGHSYFNVLVLITNVGAGADVHSYIKGSRGQMSMSHNMG	180
Db	154	VPVSEFRVPCMKKCGIFRTINGHSYFNVLVITNVGAGADVHSYIKGSRGQMSMSHNMG	213
QY	181	QNMGSNNYTLNGGSLSPQVYTSDDGSTRVATNLYVPSNMQFGQTDEBPQR	227
Db	214	QNMGSNNYTLNGQPLSPQVYVTSDDGSTRVATNLYVAPSNMQFGQTDEBPQR	260

ID	Q9SBT1	PRELIMINARY:	PRT:	253 AA.
AC	Q9SBT1.			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	EXPANSIN.			
GN	EXP2.			
OS	Fragaria ananassa (Strawberry).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;			
OX	NCBI_TaxID=3747;			
RM	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CV. CHANDLER;			
RA	Civello P.M., Sabehat A., Powell A.L.T., Bennett A.B.;			
RT	"An expansin gene expressed in ripening strawberry fruit is auxin-			
RT	independent.";			
RL	Plant Physiol. 12:1273-1279(1999).			
DR	EMBL; AF159563; AAF21101.1; -			
DR	InterPro: IPR000882; Pollen_allergen.			
DR	InterPro: IPR000408; RCCL			
DR	Pfam; PF01357; Pollen_allergen: 1.			
DR	PRINTS; PRO1225; EXPANSINFAMLY.			
DR	ProDom; PD002179; Pollen_allergen: 1.			
DR	PROSITE; PS00626; RCCL_2; UNKNOWN1.			
SO	SEQUENCE 253 AA; 26887 MW; CE339CF00ADE1CFE CRC64;			

Query Match	86.7%;	Score 1105.5;	DB 10;	Length 253;
Best Local Similarity	85.5%;	Pred. No. 7.1e-94;		
Matches 195;	Conservative 14;	Mismatches 18;	Indels 1;	Gaps 1;

```

QY      1  DYG--GMOGHAHTPEYGGGASGTMGACGYGMLYSGQGTNTVALSTALFPNNGLSGCGFE  59
Db      26  DYGAGWAGGHAHTPEYGGGASGTMGACGYGMLYSGQGTNTVALSTALFPNDGLSCGSCYE  85

QY      60  MTCNTDPKMCLEGTIRVATATNEFCPPNFALPNDGGMCPRLQHFDMAEPAFLQIAOYRAG  113
Db      86  MRCDDNPDMCPLPGSTIIVATNEFCPPNFQAQNDNGCMCPRLQHFDLAEPAFLQIAOYRAG  143

QY      120  IVPVSFRVPCMKKGGVAFETINGHSYFNLYLITVAGGAGDVHSYIKGSRFGQMSMSRNW  173
Db      146  IVPVSFRVAVCVKKGGRIFETINGHSYFNLYLITVAGGAGDVHSYIKGSKGQMSMSRNW  203

QY      180  GONNOSNNYTLNGGGSFPOVTLSDRRTLEAYLVPNNMQFGOTYEGPQF  227
Db      206  GONNOSNNYTLNGQALSFQVTTISDGRITYSNNVAPGNMQFGOTYEGGQF  253

```

RESULT 7			
093XP2			
ID	093XP2	PRELIMINARY;	PRT; 253 AA.
AC	093XP2		
DT	01-DEC-2001 (TREMBLrel. 19, Created)		
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)		
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)		
DE	EXPANSIN.		
GN	EXP1.		
OS	Prunus cerasus.		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;		
OC	eucrosid_1; Rosales; Rosaceae; Amygdaloideae; Prunus.		
OX	NCBI_TaxID=140311;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=RIPENING FRUIT;		
RA	Yoo S.-D., Gao Z., Cantliff C., Loeschner W., van Nocker S.;		
RT	"Coordinated expression of genes encoding expansins and other cell		
RT	wall-modifying enzymes is associated with pectin-related changes in		
RT	the cell wall during ripening of cherry ('P. cerasus') fruit.";		
RL	Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.		
EMBL	AF350936; AAK48845.1; -		

Best Local Similarity 82.7%; Pred. No. 1,1e-92;
Matches 186; Conservative 22; Mismatches 17; Indels 0; Gaps 0;

OY 2 YGWMOSGHAATFYGGDSAGTGMGACGYGNLYSGYGTNTVALSTALFNNGLSGACACEMT 61
DB 7 YGWMESAHATFYGGSDASGTGMGACGYGNLYSGYGTNTVALSTALFNNGLSGACACEMR 66
OY 62 CTNDPKMCLPCTIRVTATNFCPPNFALPNDGCMCPNPLQHPDMAEPALQIAQYRAGIV 121
DB 67 CNDPQWMLPGVTYVTAATNFCPPNFALPNDGCMCPNPLQHPDMAEPALQIAQYRAGIV 126
OY 122 PVSFRRVPCMKKGVFTTNGHSYFNVLITNVGAGDVHSVSIKSGRTGMSMRNMQ 181
DB 127 PLYTRVPCLRKGGIRFTVNGHSYFNVLITNVGAGDVHVAIVSIKSGRSGMQPMSRMNQ 186
OY 182 NMOSNNYLNQGLSFQVTLSDGRTLTAVNLVPSNMOPGQTEGQ 226
DB 187 NMOSNNYLNQGLSFQVTLSDGRTVSNVAVPSNMOPGQTEGQ 231

RESULT 11

O9SMY1

ID O9SMY1 PRELIMINARY; PRT: 253 AA.

AC O9SMY1: 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE EXPANSIN.
OS Pinus taeda (loblolly pine).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
OX NCBI_TaxID=3352;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE-HYPOCOTYL;
RA MEDLINE=99329318; PubMed=10398718;
RA Hutchison K.W., Singer P.B., McIntosh S., Diaz-Sala C., Greenwood M.S.;
RT "Expansins are conserved in conifers and expressed in hypocotyls in
RT response to exogenous auxin."
RL Plant Physiol. 120:827-832(1999).
DR EMBL: AF085330; AAD47901.1;
DR InterPro: IPR000882; Pollen_allergen.
DR Pfam: PF01357; Pollen_allergen; 1.
DR PRINTS: PR01225; EXPANSINFAMILY.
DR Prodom: PD002179; Pollen_allergen; 1.
SQ SEQUENCE 253 AA: 27150 MW: 2675BBI569BD75D4 CRC64;

Query Match

Best Local Similarity 85.6%; Score 1092; DB 10; Length 253;
Matches 186; Conservative 22; Mismatches 17; Indels 0; Gaps 0;

OY 2 YGWMOSGHAATFYGGDSAGTGMGACGYGNLYSGYGTNTVALSTALFNNGLSGACACEMT 61
DB 28 YGWMESAHATFYGGSDASGTGMGACGYGNLYSGYGTNTVALSTALFNNGLSGACACEMR 87
OY 62 CTNDPKMCLPCTIRVTATNFCPPNFALPNDGCMCPNPLQHPDMAEPALQIAQYRAGIV 121
DB 88 CNDPQWMLPGVTYVTAATNFCPPNFALPNDGCMCPNPLQHPDMAEPALQIAQYRAGIV 147
OY 122 PVSFRRVPCMKKGVFTTNGHSYFNVLITNVGAGDVHSVSIKSGRTGMSMRNMQ 181
DB 148 PLYTRVPCLRKGGIRFTVNGHSYFNVLITNVGAGDVHVAIVSIKSGRSGMQPMSRMNQ 207
OY 182 NMOSNNYLNQGLSFQVTLSDGRTLTAVNLVPSNMOPGQTEGQ 226
DB 208 NMOSNNYLNQGLSFQVTLSDGRTVSNVAVPSNMOPGQTEGQ 252

RESULT 12

P93492

ID P93492 PRELIMINARY; PRT: 232 AA.

AC P93492: 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Created)

DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE EXPANSIN (FRAGMENT).
OS Pinus taeda (loblolly pine).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
OX NCBI_TaxID=3352;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE-HYPOCOTYL;
RA Hutchison K.W., Singer P.B., Diaz-Sala C., Greenwood M.S.;
RT "Expansins are conserved in conifers and expressed in response to
RT exogenous auxin."
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: U64890; AAB40634.1;
DR InterPro: IPR000882; Pollen_allergen.
DR Pfam: PF01357; Pollen_allergen; 1.
DR PRINTS: PR01225; EXPANSINFAMILY.
DR Prodom: PD002179; Pollen_allergen; 1.
FT NON_TER 1
SQ SEQUENCE 232 AA: 24734 MW: 1340AF483DE0741F CRC64;

Query Match

Best Local Similarity 85.3%; Score 1087; DB 10; Length 232;
Matches 185; Conservative 22; Mismatches 18; Indels 0; Gaps 0;

OY 2 YGWMOSGHAATFYGGDSAGTGMGACGYGNLYSGYGTNTVALSTALFNNGLSGACACEMT 61
DB 7 YGWMESAHATFYGGSDASGTGMGACGYGNLYSGYGTNTVALSTALFNNGLSGACACEMR 66
OY 62 CTNDPKMCLPCTIRVTATNFCPPNFALPNDGCMCPNPLQHPDMAEPALQIAQYRAGIV 121
DB 67 CNDPQWMLPGVTYVTAATNFCPPNFALPNDGCMCPNPLQHPDMAEPALQIAQYRAGIV 126
OY 122 PVSFRRVPCMKKGVFTTNGHSYFNVLITNVGAGDVHSVSIKSGRTGMSMRNMQ 181
DB 127 PLYTRVPCLRKGGIRFTVNGHSYFNVLITNVGAGDVHVAIVSIKSGRSGMQPMSRMNQ 186
OY 182 NMOSNNYLNQGLSFQVTLSDGRTLTAVNLVPSNMOPGQTEGQ 226
DB 187 NMOSNNYLNQGLSFQVTLSDGRTVSNVAVPSNMOPGQTEGQ 231

RESULT 13

P93495

ID P93495 PRELIMINARY; PRT: 232 AA.

AC P93495: 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE EXPANSIN (FRAGMENT).
OS Pinus taeda (loblolly pine).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
OX NCBI_TaxID=3352;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE-HYPOCOTYL;
RA Hutchison K.W., Singer P.B., Diaz-Sala C., Greenwood M.S.;
RT "Expansins are conserved in conifers and expressed in response to
RT exogenous auxin."
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: U64893; AAB40637.1;
DR InterPro: IPR000882; Pollen_allergen.
DR Pfam: PF01357; Pollen_allergen; 1.
DR PRINTS: PR01225; EXPANSINFAMILY.
DR Prodom: PD002179; Pollen_allergen; 1.
FT NON_TER 1
SQ SEQUENCE 232 AA: 24694 MW: 9238EACD9E1F6C5E CRC64;

Query Match

Best Local Similarity 85.1%; Score 1085; DB 10; Length 232;
Matches 185; Conservative 23; Mismatches 17; Indels 0; Gaps 0;

QY 2 YGWSGHATFFYGGGASGTMGACGYGNLYSOGYGTNTVALSTALFNNGLSCGACFE 61
 DB 7 YGWSGHATFFYGGGASGTMGACGYGNLYSOGYGTNTVALSTALFNNGLSCGACFE 66
 QY 62 CTNDPKWCLPCTITVATNFCPPNFALPNDGWCNPPLOHFMARPAFLQIAQYRAGI 121
 DB 67 CNDPOMCLPCTITVATNFCPPNFALPNDGWCNPPLOHFMARPAFLQIAQYRAGI 126
 QY 122 PVSEFRVPCMKKGGVFTINGHSYFNLYLTNNVGAGDVHSVSIKSGRTGMSRNMWQ 181
 DB 127 PLTVRVCPLKGGIRFTVNGHSYFNLYLTNNVGAGDVHSVSIKSGRTGMSRNMWQ 186
 QY 182 NMOSNNLYNGGSLSFQVTLSDGRTLTAYNLVPSNMWFGQTEGPO 226
 DB 187 NMOSNNLYNGGSLSFQVTLSDGRTLTAYNLVPSNMWFGQTEGPO 231

RESULT 14
 ID 082625 PRELIMINARY; PRT; 247 AA.
 AC 082625;
 DT 01-NOV-1998 (Tremblrel. 08, Created)
 DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE EXPANSIN.
 GN LEEXP2.
 OS Lycopersicon esculentum (tomato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 OC NCBI_TaxID=4081;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=HYPOCOTYL;
 RA Catala C., Rose J.K.C., Bennett A.B.;
 RT "Hormonal regulation and expression patterns of LeEXP2, a new tomato
 RT expansin.";
 RT Submitted (Oct-1998) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. VEN8;
 RA Cadevas D., Muster M., Vogler H., Mandel T., Rose J.,
 RA MCoenen-Mason S., Kuhlmeier C.;
 RT "limited correlation between expansin gene expression and elongation
 RT growth rate.";
 RT Submitted (May-1999) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP EMBL; AF096776; AAC64201.1;
 DR EMBL; AJ239068; CAB43197.1;
 DR InterPro: IPR000882; Pollen_allergen.
 DR Pfam: PF01357; Pollen_allergen.1.
 DR PRINTS: PR01225; EXPANSINFAMLY.
 DR Prodom: PD002179; Pollen_allergen.1.
 SO SEQUENCE 247 AA; 26470 MW; 99E6F53F84F7A3E1 CRC64;

Query Match 85.0%; Score 1084; DB 10; Length 247;
 Best Local Similarity 81.9%; Pred. No. 6, 6e-92;
 Matches 186; Conservative 23; Mismatches 18; Indels 0; Gaps 0;

QY 1 DYGGWOSGHATFFYGGGASGTMGACGYGNLYSOGYGTNTVALSTALFNNGLSCGACFE 60
 DB 21 DYGGWOSGHATFFYGGGASGTMGACGYGNLYSOGYGTNTVALSTALFNNGLSCGACFE 80
 QY 61 TCTNDPKWCLPCTITVATNFCPPNFALPNDGWCNPPLOHFMARPAFLQIAQYRAGI 120
 DB 81 TCTNDPKWCLPCTITVATNFCPPNFALPNDGWCNPPLOHFMARPAFLQIAQYRAGI 140
 QY 121 PVSEFRVPCMKKGGVFTINGHSYFNLYLTNNVGAGDVHSVSIKSGRTGMSRNMWQ 180
 DB 141 PVSEFRVPCMKKGGVFTINGHSYFNLYLTNNVGAGDVHSVSIKSGRTGMSRNMWQ 200
 QY 181 NMOSNNLYNGGSLSFQVTLSDGRTLTAYNLVPSNMWFGQTEGPO 227
 DB 187 NMOSNNLYNGGSLSFQVTLSDGRTLTAYNLVPSNMWFGQTEGPO 231

DB 201 QNMOSNNLYNGGSLSFQVTLSDGRTLTAYNLVPSNMWFGQTEGPO 247

RESULT 15
 ID 095WD4 PRELIMINARY; PRT; 253 AA.
 AC 095WD4;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE EXPANSIN.
 GN Rumex palustris.
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Caryophyllidae; Caryophyllales; Polygonaceae; Rumex.
 OC NCBI_TaxID=50298;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LEAF;
 RA MEDLINE=20330988; PubMed=10872228;
 RA Vriezen W.H., De Graaf B., Mariani C., Voesebeck L.A.C.J.;
 RT "Submergence induces expansin gene expression in flooding-tolerant
 RT Rumex palustris and not in flooding-intolerant R. acetosa.";
 RT Planta 210:956-963(2000)
 DR EMBL; AF167360; AAD49956.1;
 DR InterPro: IPR000882; Pollen_allergen.
 DR Pfam: PF01357; Pollen_allergen.1.
 DR PRINTS: PR01225; EXPANSINFAMLY.
 DR Prodom: PD002179; Pollen_allergen.1.
 SO SEQUENCE 253 AA; 27063 MW; 7DC75610C80B23A8 CRC64;

Query Match 84.8%; Score 1081.5; DB 10; Length 253;
 Best Local Similarity 82.9%; Pred. No. 1, 1e-91;
 Matches 189; Conservative 17; Mismatches 21; Indels 1; Gaps 1;

QY 1 DYGGWOSGHATFFYGGGASGTMGACGYGNLYSOGYGTNTVALSTALFNNGLSCGACFE 59
 DB 26 DYGGWOSGHATFFYGGGASGTMGACGYGNLYSOGYGTNTVALSTALFNNGLSCGACFE 85
 QY 60 MTCTNDPKWCLPCTITVATNFCPPNFALPNDGWCNPPLOHFMARPAFLQIAQYRAGI 119
 DB 86 MTCTNDPKWCLPCTITVATNFCPPNFALPNDGWCNPPLOHFMARPAFLQIAQYRAGI 145
 QY 120 IYVSEFRVPCMKKGGVFTINGHSYFNLYLTNNVGAGDVHSVSIKSGRTGMSRNMWQ 179
 DB 146 IYVSEFRVPCMKKGGVFTINGHSYFNLYLTNNVGAGDVHSVSIKSGRTGMSRNMWQ 205
 QY 180 QNMOSNNLYNGGSLSFQVTLSDGRTLTAYNLVPSNMWFGQTEGPO 227
 DB 206 QNMOSNNLYNGGSLSFQVTLSDGRTLTAYNLVPSNMWFGQTEGPO 253

Search completed: October 11, 2002, 14:59:43
 Job time : 20.6007 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 11, 2002, 14:50:30 : Search time 5.86347 Seconds
(without alignments)
1499.000 Million cell updates/sec

Title: US-09-896-301-7

Perfect score: 1275
Sequence: 1 DYGGWQSGHATFYGGDASG.....AYNLVPSNMQFGQTECPGR 227

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	231.5	18.2	263	1	MPPL_PHLPR
2	229.5	18.0	265	1	MPPL_HOLLA
3	227.5	17.8	263	1	MPPL_LOLPR
4	225	17.6	246	1	MPCL_CINDA
5	216.5	17.0	269	1	MPAL_PHAQA
6	192.5	15.1	263	1	MPOL_ORYSA
7	185	14.5	191	1	MPZ1_MAIZE
8	106.5	8.4	242	1	GUN5_TRIRE
9	100.5	7.9	316	1	Y007_CHLTR
10	97.5	7.6	181	1	GUN_MYRED
11	96	7.5	333	1	CATJ_MOUSE
12	95.5	7.5	2871	1	FBN1_MOUSE
13	93.5	7.3	2871	1	FBN1_BOVIN
14	92	7.2	1385	1	YMS5_CAEEL
15	89.5	7.0	1429	1	L112_CAEEL
16	89	7.0	677	1	SP87_DICDI
17	88.5	6.9	894	1	ILR3_HUMAN
18	87.5	6.9	316	1	Y275_CHLMU
19	87.5	6.9	2871	1	FBN1_HUMAN
20	87.5	6.9	2871	1	FBN1_PIG
21	85	6.7	154	1	KRSC_CHICK
22	84	6.6	386	1	KR87_DROME
23	83.5	6.5	633	1	C2AD_BACTU
24	83	6.5	82	1	KRHA_SHEEP
25	82	6.4	207	1	EGG2_SCHUA
26	82	6.4	212	1	EGG1_SCHUA
27	82	6.4	645	1	K22E_HUMAN
28	79.5	6.2	684	1	FBL1_CHICK
29	79	6.2	493	1	CHT1_RHINI
30	79	6.2	2005	1	CIN2_HUMAN
31	79	6.2	5263	1	FBOH_BOVMO
32	78.5	6.2	283	1	POR3_BOVIN
33	78.5	6.2	283	1	POR3_RABIT

34	78	6.1	571	1	FLA1_CAMEE	P56963 campylobact
35	77.5	6.1	574	1	FLA3_CAMEE	O46113 campylobact
36	77	6.0	465	1	GRP2_PHAVU	P10496 phaseolus v
37	77	6.0	1325	1	YDEK_ECOLI	P32051 escherichia
38	76.5	6.0	283	1	POR3_HUMAN	O92277 homo sapien
39	76.5	6.0	283	1	POR3_MOUSE	O60931 mus musculu
40	76.5	6.0	489	1	ABF3_ASPNG	P42255 aspergillus
41	76.5	6.0	680	1	PBP2_STRPN	P10524 streptococc
42	76.5	6.0	1456	1	MANR_HUMAN	P22897 homo sapien
43	76	6.0	373	1	ADHX_RABIT	O19053 oryctolagus
44	76	6.0	385	1	PER_DROME	P1686 drosophila
45	76	6.0	593	1	KICD_HUMAN	P13645 homo sapien

ALIGNMENTS

RESULT 1

ID	MPPL_PHLPR	STANDARD:	PRT:	263 AA.
AC	P43213:			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DE	01-MAR-2002 (Rel. 41, Last annotation update)			
DE	Pollen allergen Phl p 1 precursor (Phl p I).			
GN	PHLP.			
OS	Phleum pratense (Common timothy).			
OC	Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:			
OC	Spermatophyta: Magnoliophyta: Liliopsida: Poales: Poaceae: Pooidae:			
OC	Poaceae: Phleum.			
OX	NCBI_Taxid=15957;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Pollen;			
RX	MEDLINE=95015525; PubMed=7930302;			
RA	Laffer S., Valente R., Vrtala S., Susani M., van Ree R., Kraft D.,			
RA	Schneider O., Duchene M.;			
RT	"Complementary DNA cloning of the major allergen Phl p I from timothy			
RT	grass (Phleum pratense); recombinant Phl p I inhibits IgE binding to			
RT	group I allergens from eight different grass species";			
RL	J. Allergy Clin. Immunol. 94:689-698(1994).			
CC	-1- SUBCELLULAR LOCATION: Secreted.			
CC	-1- SIMILARITY: BELONGS TO THE LOI p I FAMILY OF ALLEGENS.			
CC	-1- SIMILARITY: CONTAINS 1 EXPANSTIN-LIKE EG45 DOMAIN.			
CC	-1- SIMILARITY: CONTAINS 1 EXPANSTIN-LIKE CBD DOMAIN.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/			
CC	or send an email to license@sib-sib.ch).			
DR	EMBL: X78813; CAAS5390.1; .			
DR	HSSP: P43214; IWHO.			
DR	InterPro: IPR000882; Pollen.allergen.			
DR	Pfam: PF01357; Pollen.allergen.1.			
DR	PRINTS: PR01225; EXPANSTIN-FAMILY.			
DR	ProDom: P0002179; Pollen.allergen.1.			
DR	PROSITE: PS50843; EXPANSTIN_CBD; 1.			
DR	PROSITE: PS50842; EXPANSTIN_EG45; 1.			
KW	Allergen: Glycoprotein; Signal: Multigene family.			
FT	SIGNAL	1	23	POTENTIAL.
FT	CHAIN	24	263	POLLEN ALLERGEN PHL P 1.
FT	DOMAIN	61	167	EXPANSTIN-LIKE EG45.
FT	DOMAIN	181	262	EXPANSTIN-LIKE CBD.
FT	CARBOHYD	32	32	N-LINKED (GLCNAC. . .) (POTENTIAL).
SO	SEQUENCE	263 AA;	28457 MW;	O46BA249C17BC048 CRC64;

Query Match 18.2% Score 231.5; DB 1; Length 263;
Best Local Similarity 27.6% Pred. No. 6.2e-13;
Matches 67; Conservative 41; Mismatches 96; Indels 39; Gaps 11;

QY 2 YGG-WOSGHATFYGGDASGTM--GGACGYGNLYSOGYGTNTVALSTALFNNGISCGACF 58
 DB 37 YGDKMLDAKSTWYKPTGAKRDNAGAGYKDVDPKPPSGMTGCGNTPIFKSGGSCGF 96
 QY 59 EMTCTNDPKWCL--PGTTRVATNFCPPNFALPNDGWCNPP--QHFDMAEPFLQIA 114
 DB 97 EIKCTK-PEACSGEPVYVHTDDN-----EEPIAPYHFDLSGHAFGMA 139
 QY 115 Q-----YRAGIVPSFRPVCKMKKGGVREPTIN-----GHSYFNLVLTITNGAGADVSHV 163
 DB 140 KKGEGQKTRSGELELQFRKRYCKTPEGTKTFHVEKGSNNYIALLVKYVNGDGVAV 199
 QY 164 SIK-GSRTGWSMSRNNQGNQ--SNNYLNGGSLFOVTLSDGRTLTAIVNLPVSNMFGQ 220
 DB 200 DIKKGKDKMIEKESWGAVWRVDPKLTGP-FTVRYTTEGKTAEADVIEGKADT 258
 QY 221 TYE 223
 DB 259 SYE 261

RESULT 2
 MPIL_HOLLA STANDARD; PRT; 265 AA.
 ID MPIL_HOLLA
 AC P43216; Q39975;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Major pollen allergen Hol 1 precursor (Hol 1 I) (Hol 1 I.0101 and 1.0102).
 OS Holcus lanatus (Velvet grass).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Poaceae; Holcus.
 NC NCBL_TaxID=29679;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. AVENEA; TISSUE-Pollen;
 RA Schramm G.D., Bufe A., Becker W.M., Schlaak M.;
 RL Submitted (NOV-1993) to the EMBL/GenBank/DBD databases.
 RN [2]
 RP SEQUENCE OF 18-265 FROM N.A.
 RC STRAIN-CV. AVENEA; TISSUE-Pollen;
 RA MEDLINE-97358126; PubMed-9215246;
 RT "Schramm G.D., Bufe A., Petersen A., Haas H., Schlaak M., Becker W.M.;
 RT "Mapping of IgE-binding epitopes on the recombinant major group I
 RT allergen of velvet grass pollen, Hol 1 I."
 RL J. Allergy Clin. Immunol. 99:781-787(1997).
 RN [3]
 RP CHARACTERIZATION.
 RC STRAIN-CV. AVENEA; TISSUE-Pollen;
 RA MEDLINE-96319506; PubMed-8768803;
 RT "Schramm G.D., Petersen A., Bufe A., Schlaak M., Becker W.M.;
 RT "Identification and characterization of the major allergens of velvet
 RT grass (Holcus lanatus), Hol 1 I and Hol 1 5."
 RL Int. Arch. Allergy Immunol. 110:354-363(1996).
 CC [1]- SUBCELLULAR LOCATION: Secreted.
 CC [1]- SIMILARITY: BELONGS TO THE LIL P I FAMILY OF ALLERGENS.
 CC [1]- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE EG45 DOMAIN.
 CC [1]- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE CBD DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: Z27084; CA81610.1; -;
 DR EMBL: Z68893; CA93121.1; -;
 DR HSSP: P43214; IWHO.

DR InterPro: IPR000882; Pollen.allergen.
 DR Pfam: PF01357; Pollen.allergen.1.
 DR PRINTS: PR01225; EXPANSINFAM1.
 DR ProDom: PD002179; Pollen.allergen.1.
 DR ProSITE: PS50843; EXPANSIN_CBD; 1.
 DR ProSITE: PS50842; EXPANSIN_EG45; 1.
 KW Allergen; signal.
 FT SIGNAL. 1 25
 FT CHAIN. 26 265
 FT DOMAIN. 63 169
 FT DOMAIN. 103 264
 FT VARIANT. 103 103
 FT SEQUENCE. 265 AA; 28590 MW; 1FE23B3BE198AD6D CRC64;
 Query Match 18.0%; Score 229.5; DB 1; Length 265;
 Best Local Similarity 27.6%; Pred. No. 9.3e-13;
 Matches 67; Conservative 41; Mismatches 96; Indels 39; Gaps 11;
 QY 2 YGG-WOSGHATFYGGDASGTM--GGACGYGNLYSOGYGTNTVALSTALFNNGISCGACF 58
 DB 39 YGDEMILDAKSTWYKPTGAKRDNAGAGYKDVDPKPPSGMTGCGNTPIFKSGGSCGF 98
 QY 59 EMTCTNDPKWCL--PGTTRVATNFCPPNFALPNDGWCNPP--QHFDMAEPFLQIA 114
 DB 99 EIKCTK-PEACSGEPVYVHTDDN-----EEPIAPYHFDLSGHAFGMA 141
 QY 115 Q-----YRAGIVPSFRPVCKMKKGGVREPTIN-----GHSYFNLVLTITNGAGADVSHV 163
 DB 142 KKGEGQKTRSGELELQFRKRYCKTPEGTKTFHVEKGSNNYIALLVKYVNGDGVAV 201
 QY 164 SIK-GSRTGWSMSRNNQGNQ--SNNYLNGGSLFOVTLSDGRTLTAIVNLPVSNMFGQ 220
 DB 202 DIKKGKDKMIEKESWGAVWRVDPKLTGP-FTVRYTTEGKTAEADVIEGKADT 260
 QY 221 TYE 223
 DB 261 SYE 263

RESULT 3
 MPIL_LOLPR STANDARD; PRT; 263 AA.
 ID MPIL_LOLPR
 AC P14946; P19964;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Pollen allergen Lol p I precursor (Lol p I) (Allergen R7).
 OS Lolium perenne (perennial ryegrass).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Poaceae; Lolium.
 NC NCBL_TaxID=4522;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE-90375479; PubMed-1697854;
 RA Perez M., Ishoka G.Y., Walker L.E., Chesnut R.W.;
 RT "cDNA cloning and immunological characterization of the rye grass
 RT allergen Lol p I."
 RL J. Biol. Chem. 265:16210-16215(1990).
 RN [2]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 24-48.
 RC TISSUE-Pollen;
 RA MEDLINE-91160716; PubMed-2001733;
 RA Griffith I.J., Smith P.M., Pollock J., Theerakulpisut P.,
 RA Avjoglu A., Davies S., Hough T., Singh M.B., Simpson R.J., Ward L.D.,
 RA Knox R.B.;
 RT "Cloning and sequencing of Lol pI, the major allergenic protein of
 RT rye-grass pollen."
 RL FEBS Lett. 279:210-215(1991).
 RN [3]
 RP SEQUENCE OF 24-53.
 RC TISSUE-Pollen;
 RA MEDLINE-8624068; PubMed-3718469;

RA Cottam G.P., Moran D.M., Standing R.;
RT "Phytochemical and immunochemical characterization of allergenic
RT proteins from rye-grass (Lolium perenne) pollen prepared by a rapid
RT and efficient purification method.";
RL Biochem. J. 234:305-310(1986).
RN (4)
RP SEQUENCE OF 236-263.
RX MEDLINE-89364850; PubMed-2475768;
RA Esch R.E., Klapper D.G.;
RT "Isolation and characterization of a major cross-reactive grass group
RT I allergenic determinant";
RL Allergenic determinant";
RL Mol. Immunol. 26:557-561(1989).
CC -1- SUBCELLULAR LOCATION: secreted.
CC -1- DISEASE: CAUSES GRASS POLLEN ALLERGY.
CC -1- SIMILARITY: BELONGS TO THE LOL P I FAMILY OF ALLERGENS.
CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE EG45 DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE CBD DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: M57474; AAA63279.1; -
DR EMBL: M57476; AAA63278.1; -
DR PIR: A23341; A23341.
DR PIR: B37881; B37881.
DR PIR: S13614; S13614.
DR HSSP: P43214; 1WHO.
DR InterPro: IPR000882; Pollen_allergen.
DR Pfam: PF01357; Pollen_allergen. 1.
DR PRINTS: PR01225; EXPANSIN-FAMILY.
DR ProDom: PD002179; Pollen_allergen. 1.
DR PROSITE: PS50843; EXPANSIN_CBD. 1.
DR PROSITE: PS50842; EXPANSIN_EG45. 1.
KW Allergen; Glycoprotein; Signal; Multigene family.
FT SIGNAL 1 23
FT CHAIN 1 23
FT DOMAIN 24 263
FT DOMAIN 61 167
FT DOMAIN 181 262
FT CARBOHYD 32 32
FT VARIANT 68 68
FT VARIANT 177 177
FT VARIANT 210 210
FT VARIANT 246 246
FT CONFLICT 28 28
FT CONFLICT 31 31
FT CONFLICT 48 48
SQ SEQUENCE 263 AA; 28438 MW; 7675896F279C88C9 CRC64;

Query Match 17.8%; Score 227.5; DB 1; Length 263;
Best Local Similarity 28.2%; Pred. No. 1.4e-12;
Matches 68; Conservative 38; Mismatches 100; Indels 35; Gaps 11;

QY 1 DYGG-MQSGHATFYGGGDAAGM--GGAGCGNLYSGGYGTNTVALSTLAFNNGLSGAGC 57
DB 36 EYDGLKLDKASTYGRKPTGAGPKDNGGACGYKAVDKAPFGMGCTGCTPFKGRGCGSC 95
QY 58 FEMTCTNDPKMCL-PGTRVATNFCPPNFALPNDGCMCNPL--OHFMAEPAPLQIAO 115
DB 96 FEIKCKR-PESGGEAVTVIT-----DDN-----EETAPRPHFDLSGAFSGMAK 140
QY 116 -----YRAGIVPVSFRVPCMKKGGVARTTNGHSTFN---LVLTNNGAGDVHVS 164
DB 141 KGEQONVRSGEELQFRVCKYRPDDTKPTFVERKASNNYLAIIKVVYDGGDGVAVD 200
QY 165 IK-GSRTGQMSRMNGMNOQ--SNVYLNQGLSFOVTLSDGRTLTAVNLVPSNMQFGOT 221
DB 201 IKKGKDKMTELKSGKAVRITDPLKLTGP-FTVRYTTEGKTKSEPDVIFEGKADTS 259

QY 222 Y 222
DB 260 Y 260

RESULT 4
MOTIF_CYCNA STANDARD: PRT; 246 AA.
AC 004701;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Major pollen allergen Cyn d 1.
GN CYND1.
OS Cynodon dactylon (Bermuda grass).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Chloridaceae; Cydonaceae; Cydonon.
OX NCBI_TaxID=28909;
[1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=pollen.
RX MEDLINE-96347957; PubMed-8757211;
RA Smith P.M., Suphloglu C., Griffith I.J., Theriault K., Knox R.B.,
RA Singh M.B.;
RT "Cloning and expression in yeast Pichia pastoris of a biologically
RT active form of Cyn d 1, the major allergen of Bermuda grass pollen.";
RL J. Allergy Clin. Immunol. 98:331-343(1996).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DISEASE: CAUSES GRASS POLLEN ALLERGY.
CC -1- SIMILARITY: BELONGS TO THE LOL P I FAMILY OF ALLERGENS.
CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE EG45 DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE CBD DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: S83343; IAB50734.2; -
DR HSSP: P43214; 1WHO.
DR InterPro: IPR000882; Pollen_allergen.
DR Pfam: PF01357; Pollen_allergen. 1.
DR PRINTS: PR01225; EXPANSIN-FAMILY.
DR ProDom: PD002179; Pollen_allergen. 1.
DR PROSITE: PS50843; EXPANSIN_CBD. 1.
DR PROSITE: PS50842; EXPANSIN_EG45. 1.
KW Allergen.
FT DOMAIN 39 145
FT CARBOHYD 9 240
FT DOMAIN 159 240
SQ SEQUENCE 246 AA; 26888 MW; 43D8442DBA583322 CRC64;

Query Match 17.6%; Score 225; DB 1; Length 246;
Best Local Similarity 27.1%; Pred. No. 2.1e-12;
Matches 65; Conservative 44; Mismatches 97; Indels 34; Gaps 10;

QY 2 YGG-MQSGHATFYGG--GDASGTMGAGCGYGNLYSGGYGTNTVALSTLAFNNGLSGAGC 57
DB 14 YGSKMLEARATFYGSPRCAAPDDHGGACGYKDVDPDPCMTAGCNEPTFKDGLGCRAC 73
QY 58 FEMTCTNDPKMCL-PGTRVATNFCPPNFALPNDGCMCNPL--OHFMAEPAPLQIAO 115
DB 74 YEIKC-KEPVECSGEPVLTAKITDKNY--EHLA-----AYHFDLSGKAFSGAMAK 118
QY 116 -----YRAGIVPVSFRVPCMKKGGVARTTNGHSTFN---LVLTNNGAGDVHVS 164
DB 119 KGEQONVRSGEELQFRVCKYRPDDTKPTFVERKASNNYLAIIKVVYDGGDGVAVD 178
QY 165 IKGSRTG-MQSMRMNGMNOQ--SNVYLNQGLSFOVTLSDGRTLTAVNLVPSNMQFGOT 222

[illegible][illegible]

OY 80 NECPNFPALPNDGWCNPNL--QHEDMAEPALQIAQ-----YRAGIVPSFRBPCKK 133
 DB 119 N-----DEPIAAYHFDLSGLAMKDGDEELRKAGLITDTPRRVCKYP 162
 OY 134 GGVRFIN-----GHSYFNLVLTNNVGAGADVHSYI--KSGRTGQMSRNMQMSNN 187
 DB 163 ADRTKEFHEKASNPNYLALLVYVAGDGVVEYKEKSGSE--WKALKESGALWIRIDT 221
 OY 188 YLNGOC-LSFOVTLSDGRTLTAYNLVP 213
 DB 222 PKPLKGPSVRYTTEGARSSAEDAIIP 248

RESULT 7

MP21_MAIZE
 ID MP21_MAIZE STANDARD: PRT: 191 AA.
 AC 007154:
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DE Pollen allergen Zea m 1 (Zea m 1).
 DE Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
 OC Palcoideae; Andropogoneae; Zea.
 OK NCBI_TaxID=4577;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pollen;
 RX MEDLINE=94010312; PubMed=8406014;
 RA Broadwater A.H., Rudinstein A.L., Chay C.H., Klapper D.G.,
 RA Bedinger P.A.;
 RT "Zea m1, the maize homolog of the allergen-encoding lol p1 gene of
 RT rye grass.";
 RL Gene 131:227-230(1993).
 CC -1- TISSUE SPECIFICITY: POLLEN TISSUE.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSION LOW BEFORE AND HIGH AFTER
 CC POLLEN MITOSIS.
 CC -1- DISEASE: CAUSES MAIZE POLLEN ALLERGY.
 CC -1- SIMILARITY: BELONGS TO THE LOL P 1 FAMILY OF ALLEGENS.
 CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE EG45 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE CBD DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb-sdb.ch).

DR EMBL: L14271; AAA33496.1; -
 DR PIR: JCI524; JCI524.
 DR HSSP: P43214; JCI524.
 DR MAIZE: 65840; -
 DR InterPro: IPR000882; Pollen.allergen.
 DR Pfam: PF01357; Pollen.allergen.1.
 DR PRINTS: PR01225; EXPANSIN.FAMILY.
 DR ProDom: PD002179; Pollen.allergen.1.
 DR PROSITE: PS00843; EXPANSIN.CBD.1.
 DR PROSITE: PS00842; EXPANSIN.EG45.1.
 KW Allergen; Multigene family.
 FT DOMAIN 1 91 EXPANSIN-LIKE EG45.
 FT SIGNAL 105 186 EXPANSIN-LIKE CBD.
 SO SEQUENCE 191 AA; 21362 MW; 6E2A9DF921C45C63 CRC64;

Query Match 14.5%; Score 185; DB 1; Length 191;
 Best Local Similarity 26.3%; Pred. No. 4.2e-09;
 Matches 52; Conservative 36; Mismatches 80; Indels 30; Gaps 8;

OY 40 TVALSTALNNGLSCGACFEMCTNDPKWCL--PGTIRVATNCPNFPALPNDGWCN 97

DB 2 TACGNVPIFKDGKCGSCYEVCKEKEPE-CSGNPVTEFITDMNVEP-----IAP----- 49
 OY 98 PLOHEDMAEPALQIAQ-----YRAGIVPSFRBPCKKGGVRFITNGHSYFN--- 147
 DB 50 ---YHFDLSGKAFGLAKGLNDKLRHCGIMDEYRRVRCRYPPGOKIVFITEKGCNPNY 106
 OY 148 -LVLTNNVGAGADVHSYISIKSGRTG--WQMSRNMQMSNNYLNNGOC-LSFOVTLSDGR 204
 DB 107 VAVLVKVFVADDDIVLMEIODKLSAEMKPKLSMCAIWRMDTAAALGPPSIRLTSKSG 166
 OY 205 TLTAVNLVPSNMQFGQTY 222
 DB 167 KVIAKDIIPANMRPDAVY 184

RESULT 8

GUN5_TRIE
 ID GUN5_TRIE STANDARD: PRT: 242 AA.
 AC P43317:
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DE Endoglucanase V precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase V)
 DE (Cellulase V) (EG V).
 DE EGS5.
 OS Trichoderma reesei (Hypocrea jecorina).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Hypocreales; Hypocreaceae; Hypocrea.
 OK NCBI_TaxID=51453;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=QM9414 / RUF C-30;
 RX MEDLINE=95075308; PubMed=7984103;
 RA Saohelmo A., Henrissat B., Hoffer A.-M., Teleman O., Penttilae M.,
 RA "A novel, small endoglucanase gene, eg15, from Trichoderma reesei
 RT isolated by expression in yeast.";
 RL Mol. Microbiol. 13:219-228(1994).
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
 CC linkages in cellulose.
 CC -1- SIMILARITY: BELONGS TO CELLULOSE FAMILY K (FAMILY 45 OF GLYCOSYL
 CC HYDROLASES).
 CC -1- SIMILARITY: CONTAINS 1 FUNGAL-TYPE CELLULOSE-BINDING DOMAIN (CBD).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb-sdb.ch).

DR EMBL: Z33381; CAAB3846.1; -
 DR HSSP: P00725; ZCBH.
 DR InterPro: IPR000254; CBD_fungal.
 DR InterPro: IPR000334; Glyco_hydro_45.
 DR Pfam: PF00734; CBD.1; 1.
 DR Pfam: PF02015; Glyco_hydro_45; 1.
 DR ProDom: PD001821; CBD_fungal.1.
 DR SMART: SM00236; fCBD.1.
 DR PROSITE: PS00362; CBD_FUNGAL.1.
 DR PROSITE: PS00842; EXPANSIN.EG45.1.
 DR PROSITE: PS01140; GLYCOSYL_HYDROL_F45.1.
 KW Cellulose degradation; Hydrolase; Glycosidase; Signal.
 FT CHAIN 1 17 POTENTIAL.
 FT ACT_SITE 18 242 ENDOGLUCANASE V.
 FT ACT_SITE 27 27 NUCLEOPHILE (BY SIMILARITY).
 FT DOMAIN 134 134 PROTON DONOR (BY SIMILARITY).
 FT DOMAIN 183 182 CATALYTIC.
 FT DOMAIN 206 205 PRO/SER-RICH (LINKER).
 FT CARBOHYD 182 182 CELLULOSE-BINDING (BY SIMILARITY).
 FT DISULFID 213 210 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT 224 240 BY SIMILARITY.
 FT 224 240 BY SIMILARITY.


```

Oy 31 LYSQGVGTNTVALSTALFENNGSCGACEMTCTDPKMKCLPGTIRYATNF--CPNPAI. 88
Db 175 KGCQS--GTHAQAEFVYLKKNKGLEAATPYBGKGGPCPRYSSENASANITDYVAL.PPN--- 230
Oy 89 PNDGGMCPPLQHDMAEPALFLAQAGRAIVDPV-----SPRRPCKMKGGVPT 139
Db 231 -----ELYLWAV--ASIGVSNADIASDSTR-----FYNGIAYE 265
Oy 140 INSHF--NLVLTINVGAGDVSVSIKSGRGQSMQSRNQMNQSNNYL 169
Db 266 PNCSSYFNHVALVYGYGSEGV-----KQNNYV-LIKNSMEENGMMGYM 311

RESULT 12
FBN1_BOVIN STANDARD: PRT; 2871 AA.
ID FBN1_BOVIN
AC P98133;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fibrillin 1 precursor (MP340).
OS FBN1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bovina; Bos.
OX NCBI_TaxID=9913;
RX [1]
RC SEQUENCE FROM N.A.
RP TISSUE=Skin;
RX MEDLINE=95137597; PubMed=7835900;
RX Tlstara D.J., Potler K.A., Byers P.H.;
RT "Sequence of the coding region of the bovine fibrillin cDNA and
RT localization to bovine chromosome 10.";
RL Genomics 23:480-485(1994).
RN [2]
RP PARTIAL SEQUENCE.
RX MEDLINE=96132851; PubMed=8557636;
RX Gibson M.A., Hatzinikolas G., Kumaratillake J.S., Sandberg L.B.,
RX Nicholl J.K., Sutherland G.R., Cleary E.G.;
RT "Further characterization of proteins associated with elastic fiber
RT microfibrils including the molecular cloning of MAGP-2 (MP25).";
RL J. Biol. Chem. 271:1096-1103(1996).
CC -1- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS
CC THAT BINDS CALCIUM. FIBRILLIN-1-CONTAINING MICROFIBRILS
CC LONG-TERM FORCE BEARING STRUCTURE.
CC -1- PWM: FORMS INTERMOLECULAR DISULFIDE BONDS EITHER WITH OTHER
CC FIBRILLIN-1 MOLECULES OR WITH OTHER COMPONENTS OF THE
CC MICROFIBRILS.
CC -1- SIMILARITY: CONTAINS 43 CALCIUM-BINDING AND 4 NON-CALCIUM BINDING
CC EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 7 TGF-BETA BINDING PROTEIN DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on
CC use, by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L28748; AAA74122.1; -.
DR HSSP; P35555; IAP7.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001981; EGF_Ca.
DR InterPro; IPR001438; EGF_T1.
DR InterPro; IPR002212; TB.
DR Pfam; PF00683; TB; 9.
DR Pfam; PF00683; TB; 9.
DR PRINTS; PR00010; EGF_CA. 42
DR SMART; SM00179; EGF_CA. 42

```

DR SMART: SM00001: EGF like: 4.
 DR PROSITE; PS00010; ASX_HYDROXYL; 43.
 DR PROSITE; PS00022; EGF_1; 2.
 DR PROSITE; PS01186; EGF_2; 38.
 DR PROSITE; PS01187; EGF_CA; 45.
 KW Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;
 KW Repeat: Signal: Multigene family.
 FT SIGNAL 1 27
 FT CHAIN 28 2871
 FT DOMAIN 81 112
 FT DOMAIN 115 146
 FT DOMAIN 147 178
 FT DOMAIN 246 287
 FT DOMAIN 288 329
 FT REPEAT 330 390
 FT REPEAT 392 446
 FT DOMAIN 449 489
 FT DOMAIN 490 529
 FT DOMAIN 530 571
 FT DOMAIN 572 612
 FT DOMAIN 613 653
 FT REPEAT 654 722
 FT REPEAT 723 764
 FT DOMAIN 765 806
 FT DOMAIN 807 846
 FT DOMAIN 910 951
 FT REPEAT 952 1027
 FT DOMAIN 1028 1069
 FT DOMAIN 1070 1112
 FT DOMAIN 1113 1154
 FT DOMAIN 1155 1196
 FT DOMAIN 1197 1237
 FT DOMAIN 1238 1279
 FT DOMAIN 1280 1321
 FT DOMAIN 1322 1362
 FT DOMAIN 1363 1403
 FT DOMAIN 1404 1445
 FT DOMAIN 1446 1486
 FT DOMAIN 1487 1527
 FT REPEAT 1528 1605
 FT DOMAIN 1606 1647
 FT DOMAIN 1648 1688
 FT REPEAT 1689 1765
 FT REPEAT 1766 1807
 FT DOMAIN 1808 1848
 FT DOMAIN 1849 1890
 FT DOMAIN 1891 1929
 FT DOMAIN 1930 1972
 FT DOMAIN 1973 2012
 FT DOMAIN 2013 2054
 FT REPEAT 2055 2126
 FT REPEAT 2127 2165
 FT DOMAIN 2166 2205
 FT DOMAIN 2206 2246
 FT DOMAIN 2247 2290
 FT DOMAIN 2291 2332
 FT REPEAT 2333 2401
 FT DOMAIN 2402 2443
 FT DOMAIN 2444 2484
 FT DOMAIN 2485 2523
 FT DOMAIN 2524 2566
 FT DOMAIN 2567 2606
 FT DOMAIN 2607 2647
 FT DOMAIN 2648 2687
 FT DISULFID 85 94
 FT DISULFID 89 100
 FT DISULFID 102 111
 FT DISULFID 119 129
 FT DISULFID 123 134
 FT DISULFID 136 145
 FT DISULFID 150 160
 FT DISULFID 154 166
 FT DISULFID 168 177

FT	DISULFID	250	262	BY SIMILARITY.
FT	DISULFID	257	271	BY SIMILARITY.
FT	DISULFID	273	286	BY SIMILARITY.
FT	DISULFID	292	304	BY SIMILARITY.
FT	DISULFID	299	313	BY SIMILARITY.
FT	DISULFID	315	328	BY SIMILARITY.
FT	DISULFID	453	465	BY SIMILARITY.
FT	DISULFID	460	474	BY SIMILARITY.
FT	DISULFID	476	488	BY SIMILARITY.
FT	DISULFID	494	504	BY SIMILARITY.
FT	DISULFID	499	513	BY SIMILARITY.
FT	DISULFID	515	528	BY SIMILARITY.
FT	DISULFID	534	546	BY SIMILARITY.
FT	DISULFID	541	555	BY SIMILARITY.
FT	DISULFID	557	570	BY SIMILARITY.
FT	DISULFID	576	587	BY SIMILARITY.
FT	DISULFID	582	596	BY SIMILARITY.
FT	DISULFID	598	611	BY SIMILARITY.
FT	DISULFID	617	628	BY SIMILARITY.
FT	DISULFID	623	637	BY SIMILARITY.
FT	DISULFID	639	652	BY SIMILARITY.
FT	DISULFID	727	739	BY SIMILARITY.
FT	DISULFID	734	748	BY SIMILARITY.
FT	DISULFID	750	763	BY SIMILARITY.
FT	DISULFID	769	781	BY SIMILARITY.
FT	DISULFID	776	790	BY SIMILARITY.
FT	DISULFID	792	805	BY SIMILARITY.
FT	DISULFID	811	821	BY SIMILARITY.
FT	DISULFID	816	830	BY SIMILARITY.
FT	DISULFID	832	845	BY SIMILARITY.
FT	DISULFID	914	926	BY SIMILARITY.
FT	DISULFID	921	935	BY SIMILARITY.
FT	DISULFID	937	950	BY SIMILARITY.
FT	DISULFID	1032	1044	BY SIMILARITY.
FT	DISULFID	1039	1053	BY SIMILARITY.
FT	DISULFID	1055	1068	BY SIMILARITY.
FT	DISULFID	1074	1086	BY SIMILARITY.
FT	DISULFID	1081	1095	BY SIMILARITY.
FT	DISULFID	1097	1111	BY SIMILARITY.
FT	DISULFID	1117	1129	BY SIMILARITY.
FT	DISULFID	1124	1138	BY SIMILARITY.
FT	DISULFID	1140	1153	BY SIMILARITY.
FT	DISULFID	1159	1171	BY SIMILARITY.
FT	DISULFID	1166	1180	BY SIMILARITY.
FT	DISULFID	1182	1195	BY SIMILARITY.
FT	DISULFID	1201	1212	BY SIMILARITY.
FT	DISULFID	1208	1221	BY SIMILARITY.
FT	DISULFID	1223	1236	BY SIMILARITY.
FT	DISULFID	1242	1254	BY SIMILARITY.
FT	DISULFID	1249	1263	BY SIMILARITY.
FT	DISULFID	1265	1278	BY SIMILARITY.
FT	DISULFID	1284	1296	BY SIMILARITY.
FT	DISULFID	1291	1305	BY SIMILARITY.
FT	DISULFID	1307	1320	BY SIMILARITY.
FT	DISULFID	1326	1339	BY SIMILARITY.
FT	DISULFID	1333	1348	BY SIMILARITY.
FT	DISULFID	1350	1361	BY SIMILARITY.
FT	DISULFID	1367	1380	BY SIMILARITY.
FT	DISULFID	1374	1389	BY SIMILARITY.
FT	DISULFID	1391	1402	BY SIMILARITY.
FT	DISULFID	1408	1420	BY SIMILARITY.
FT	DISULFID	1415	1429	BY SIMILARITY.
FT	DISULFID	1431	1444	BY SIMILARITY.
FT	DISULFID	1450	1461	BY SIMILARITY.
FT	DISULFID	1456	1470	BY SIMILARITY.
FT	DISULFID	1472	1485	BY SIMILARITY.
FT	DISULFID	1491	1502	BY SIMILARITY.
FT	DISULFID	1497	1511	BY SIMILARITY.
FT	DISULFID	1513	1526	BY SIMILARITY.
FT	DISULFID	1610	1622	BY SIMILARITY.

Query Match
 Best Local Similarity

7.58: Score 95.5; DB 1; Length 2871;
 26.68: Pred. No. 3.2;

FT	SIGNAL.	1	27	POTENTIAL.
FT	CHAIN	28	2871	FIBRILLIN 1.
FT	DOMAIN	81	112	EGF-LIKE 1, NON-CALCIUM BINDING
FT	DOMAIN	115	146	EGF-LIKE 2, NON-CALCIUM BINDING
FT	DOMAIN	147	178	EGF-LIKE 3, NON-CALCIUM BINDING
FT	DOMAIN	246	287	EGF-LIKE 4, CALCIUM-BINDING.
FT	DOMAIN	288	329	EGF-LIKE 5, CALCIUM-BINDING.
FT	DOMAIN	330	401	TCRBP 1.
FT	DOMAIN	402	446	PRO-RICH.
FT	DOMAIN	449	489	EGF-LIKE 6, NON-CALCIUM BINDING
FT	DOMAIN	490	521	EGF-LIKE 7, CALCIUM-BINDING.
FT	DOMAIN	530	571	EGF-LIKE 8, CALCIUM-BINDING.
FT	DOMAIN	572	612	EGF-LIKE 9, CALCIUM-BINDING.
FT	DOMAIN	613	653	EGF-LIKE 10, CALCIUM-BINDING.
FT	DOMAIN	656	721	TCRBP 2.
FT	DOMAIN	723	764	EGF-LIKE 11, CALCIUM-BINDING.
FT	DOMAIN	765	806	EGF-LIKE 12, CALCIUM-BINDING.
FT	DOMAIN	807	846	EGF-LIKE 13, CALCIUM-BINDING.
FT	DOMAIN	910	951	TCRBP 3.
FT	DOMAIN	952	1018	EGF-LIKE 14, CALCIUM-BINDING.
FT	DOMAIN	1028	1069	EGF-LIKE 15, CALCIUM-BINDING.
FT	DOMAIN	1070	1112	EGF-LIKE 16, CALCIUM-BINDING.
FT	DOMAIN	1113	1154	EGF-LIKE 17, CALCIUM-BINDING.
FT	DOMAIN	1155	1196	EGF-LIKE 18, CALCIUM-BINDING.
FT	DOMAIN	1197	1237	EGF-LIKE 19, CALCIUM-BINDING.
FT	DOMAIN	1238	1279	EGF-LIKE 20, CALCIUM-BINDING.
FT	DOMAIN	1280	1321	EGF-LIKE 21, CALCIUM-BINDING.
FT	DOMAIN	1322	1362	EGF-LIKE 22, CALCIUM-BINDING.
FT	DOMAIN	1363	1403	EGF-LIKE 23, CALCIUM-BINDING.
FT	DOMAIN	1404	1445	EGF-LIKE 24, CALCIUM-BINDING.
FT	DOMAIN	1446	1486	EGF-LIKE 25, CALCIUM-BINDING.
FT	DOMAIN	1487	1527	EGF-LIKE 26, CALCIUM-BINDING.
FT	DOMAIN	1528	1569	TCRBP 4.
FT	DOMAIN	1569	1647	EGF-LIKE 27, CALCIUM-BINDING.
FT	DOMAIN	1606	1648	EGF-LIKE 28, CALCIUM-BINDING.
FT	DOMAIN	1648	1688	EGF-LIKE 29, CALCIUM-BINDING.
FT	DOMAIN	1689	1758	TCRBP 5.
FT	DOMAIN	1766	1807	EGF-LIKE 30, CALCIUM-BINDING.
FT	DOMAIN	1808	1848	EGF-LIKE 31, CALCIUM-BINDING.
FT	DOMAIN	1849	1890	EGF-LIKE 32, CALCIUM-BINDING.
FT	DOMAIN	1891	1929	EGF-LIKE 33, CALCIUM-BINDING.
FT	DOMAIN	1930	1972	EGF-LIKE 34, CALCIUM-BINDING.
FT	DOMAIN	1973	2012	EGF-LIKE 35, CALCIUM-BINDING.
FT	DOMAIN	2013	2054	TCRBP 6.
FT	DOMAIN	2055	2121	EGF-LIKE 36, CALCIUM-BINDING.
FT	DOMAIN	2127	2165	EGF-LIKE 37, CALCIUM-BINDING.
FT	DOMAIN	2166	2205	EGF-LIKE 38, CALCIUM-BINDING.
FT	DOMAIN	2206	2245	EGF-LIKE 39, CALCIUM-BINDING.
FT	DOMAIN	2247	2290	EGF-LIKE 40, CALCIUM-BINDING.
FT	DOMAIN	2291	2332	TCRBP 7.
FT	DOMAIN	2333	2400	EGF-LIKE 41, CALCIUM-BINDING.
FT	DOMAIN	2402	2443	EGF-LIKE 42, CALCIUM-BINDING.
FT	DOMAIN	2444	2484	EGF-LIKE 43, CALCIUM-BINDING.
FT	DOMAIN	2485	2523	EGF-LIKE 44, CALCIUM-BINDING.
FT	DOMAIN	2524	2565	EGF-LIKE 45, CALCIUM-BINDING.
FT	DOMAIN	2567	2606	EGF-LIKE 46, CALCIUM-BINDING.
FT	DOMAIN	2607	2647	EGF-LIKE 47, CALCIUM-BINDING.
FT	DOMAIN	2648	2687	BY SIMILARITY.
FT	DISULEPID	85	94	BY SIMILARITY.
FT	DISULEPID	89	100	BY SIMILARITY.
FT	DISULEPID	102	111	BY SIMILARITY.
FT	DISULEPID	119	129	BY SIMILARITY.
FT	DISULEPID	123	134	BY SIMILARITY.
FT	DISULEPID	136	145	BY SIMILARITY.
FT	DISULEPID	150	160	BY SIMILARITY.
FT	DISULEPID	154	166	BY SIMILARITY.
FT	DISULEPID	168	177	BY SIMILARITY.
FT	DISULEPID	250	262	BY SIMILARITY.
FT	DISULEPID	257	271	BY SIMILARITY.
FT	DISULEPID	273	286	BY SIMILARITY.
FT	DISULEPID	292	304	BY SIMILARITY.
FT	DISULEPID	299	313	BY SIMILARITY.
FT	DISULEPID	315	358	BY SIMILARITY.
FT	DISULEPID	453	455	BY SIMILARITY.

```

FT DISULFID 460 474 BY SIMILARITY.
FT DISULFID 476 488 BY SIMILARITY.
FT DISULFID 494 504 BY SIMILARITY.
FT DISULFID 499 513 BY SIMILARITY.
FT DISULFID 515 528 BY SIMILARITY.
FT DISULFID 534 546 BY SIMILARITY.
FT DISULFID 541 555 BY SIMILARITY.
FT DISULFID 557 570 BY SIMILARITY.
FT DISULFID 576 587 BY SIMILARITY.
FT DISULFID 582 596 BY SIMILARITY.
FT DISULFID 598 611 BY SIMILARITY.
FT DISULFID 617 628 BY SIMILARITY.
FT DISULFID 623 637 BY SIMILARITY.
FT DISULFID 639 652 BY SIMILARITY.
FT DISULFID 727 739 BY SIMILARITY.
FT DISULFID 734 748 BY SIMILARITY.
FT DISULFID 750 763 BY SIMILARITY.
FT DISULFID 769 781 BY SIMILARITY.
FT DISULFID 776 790 BY SIMILARITY.
FT DISULFID 792 805 BY SIMILARITY.
FT DISULFID 811 821 BY SIMILARITY.
FT DISULFID 816 830 BY SIMILARITY.
FT DISULFID 832 845 BY SIMILARITY.
FT DISULFID 914 926 BY SIMILARITY.
FT DISULFID 921 935 BY SIMILARITY.
FT DISULFID 937 950 BY SIMILARITY.
FT DISULFID 1032 1044 BY SIMILARITY.
FT DISULFID 1039 1053 BY SIMILARITY.
FT DISULFID 1055 1068 BY SIMILARITY.
FT DISULFID 1074 1086 BY SIMILARITY.
FT DISULFID 1081 1095 BY SIMILARITY.
FT DISULFID 1097 1111 BY SIMILARITY.
FT DISULFID 1117 1129 BY SIMILARITY.
FT DISULFID 1124 1138 BY SIMILARITY.
FT DISULFID 1140 1153 BY SIMILARITY.
FT DISULFID 1159 1171 BY SIMILARITY.
FT DISULFID 1166 1180 BY SIMILARITY.
FT DISULFID 1182 1195 BY SIMILARITY.
FT DISULFID 1201 1212 BY SIMILARITY.
FT DISULFID 1208 1221 BY SIMILARITY.
FT DISULFID 1223 1236 BY SIMILARITY.
FT DISULFID 1242 1254 BY SIMILARITY.
FT DISULFID 1249 1263 BY SIMILARITY.
FT DISULFID 1265 1278 BY SIMILARITY.
FT DISULFID 1284 1296 BY SIMILARITY.
FT DISULFID 1291 1305 BY SIMILARITY.
FT DISULFID 1307 1320 BY SIMILARITY.
FT DISULFID 1326 1339 BY SIMILARITY.
FT DISULFID 1333 1348 BY SIMILARITY.
FT DISULFID 1350 1361 BY SIMILARITY.
FT DISULFID 1367 1380 BY SIMILARITY.
FT DISULFID 1374 1389 BY SIMILARITY.
FT DISULFID 1391 1402 BY SIMILARITY.
FT DISULFID 1408 1420 BY SIMILARITY.
FT DISULFID 1415 1429 BY SIMILARITY.
FT DISULFID 1431 1444 BY SIMILARITY.
FT DISULFID 1450 1461 BY SIMILARITY.
FT DISULFID 1456 1470 BY SIMILARITY.
FT DISULFID 1472 1485 BY SIMILARITY.
FT DISULFID 1491 1502 BY SIMILARITY.
FT DISULFID 1497 1511 BY SIMILARITY.
FT DISULFID 1513 1526 BY SIMILARITY.
FT DISULFID 1610 1622 BY SIMILARITY.
FT DISULFID 1617 1631 BY SIMILARITY.
FT DISULFID 1633 1646 BY SIMILARITY.
FT DISULFID 1652 1663 BY SIMILARITY.
FT DISULFID 1658 1672 BY SIMILARITY.

```

Query Match 7.3%: Score 93.5; DB 1; Length 2871;
 Best Local Similarity 26.6%; Pred. No. 4.8;
 Matches 29; Conservative 8; Mismatches 41; Indels 31; Gaps 4;

```

DB 1901 GNGTCRNTGSCNCRHGFILSHNNDICIDVDECATGNGNCRGQCVNIVGSFQCRKNE 1960
OY 51 GLSCG-----ACPEMT-CTNDPKWCLPRTIRYATNF-----CPNFALPND 91
DU 1961 GYEVAPDRGTCDVINECVLDPEKCAPGTCQNDGSRICPCPGYSIAND 2009

```

```

RESULT 14
YMS5_CAEEL STANDARD: PRT: 1385 AA.
ID YMS5_CAEEL
AC P34501;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Hypothetical 159.2 kDa protein K03H1.5 in chromosome III.
GN K03H1.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxId=6239;
XP
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2; PubMed=7906398;
RX MEDLINE=94150718;
RA Wilson R., Almscough R., Anderson K., Baynes C., Barks M.,
RA Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,
RA Fulton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Koopra A., Saunders D., Shownkeen R.,
RA Sims M., Smailton N., Smith A., Smith M., Sonhammer E., Staden K.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterston R., Watson A., Weinstock L., Wilkinson-Spratt J.,
RA Wohldman P.,
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans."
RL Nature 368:32-38(1994).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: Z29560; CAAB2664.1; -
DR PIR: S41028; S41028.
DR WormPep: K03H1.5; CE03459.
DR InterPro: IPR002909; IPT_TIG.
DR InterPro: IPR003886; Nidogen_ext.
DR InterPro: IPR000436; Sush1_SCR_CCP.
DR InterPro: IPR001846; Wvd.
DR Pfam: PF00084; sush1.1.
DR Pfam: PF00094; wvd.2.
DR SMART: SM00032; CCP.1.
DR SMART: SM00429; IPT.1.
DR SMART: SM00539; NIDO.1.
DR SMART: SM00216; WVD.1.
KW Hypothetical protein.
SQ SEQUENCE 1385 AA; 159181 MW; BDCB8F59CEA38C03 CRC64;

```

Query Match 7.2%: Score 92; DB 1; Length 1385;
 Best Local Similarity 23.0%; Pred. No. 3;
 Matches 50; Conservative 26; Mismatches 73; Indels 68; Gaps 10;

```

OY 2 YGNQSCHAFTYGGGASGTMG-GAGCYGNLYSGYGTNTVALSTALFNNNGSCGACPFM 60
DB 399 FGKQAMAGF-NGNGGTGYGLPYSGEGLMWLG-----FSNVLPGRWIR 446
OY 61 Y-----CTNDPKWCLPRTIRYATNFCCPNFALPNDGWCNPLQ-----HPDMARP 108

```

[illegible]

```
FT DISULFID 408 419 BY SIMILARITY.
FT DISULFID 413 429 BY SIMILARITY.
FT DISULFID 431 440 BY SIMILARITY.
FT DISULFID 507 518 BY SIMILARITY.
FT DISULFID 512 529 BY SIMILARITY.
FT DISULFID 531 540 BY SIMILARITY.
FT DISULFID 547 558 BY SIMILARITY.
FT DISULFID 552 567 BY SIMILARITY.
FT DISULFID 569 578 BY SIMILARITY.
FT DISULFID 586 597 BY SIMILARITY.
FT DISULFID 591 607 BY SIMILARITY.
FT DISULFID 609 618 BY SIMILARITY.
FT CARBOHYD 41 41 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 165 165 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 194 194 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 378 378 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 515 515 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 623 623 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 751 751 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 754 754 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 900 900 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1429 AA; 157115 MW; 255EDD7A62C025DB CRC64;
```

```
Query Match 7.0%; Score 89.5; DB 1; Length 1429;
Best Local Similarity 30.0%; Pred. No. 5.1;
Matches 27; Conservative 6; Mismatches 36; Indels 21; Gaps 4;
```

```
QY 13 YGGDASGTMG--GACGYGNLYSGYGTNYALSTALFNNGISCGACFEMTCTNDPKWCL 70
DB 663 FDSGDCSGKREPFSGRCRGNCACAD-----FPANGVCNQACNNECLYDGMDCCL 710
```

```
QY 71 PGTIR--VTATNCPNPFALPNDGWCNP 98
```

```
DB 711 PAVVRCPVKIREHCASRFA-----NGICDP 735
```

Search completed: October 11, 2002, 14:57:37
Job time : 7.86347 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 11, 2002, 14:53:05 : Search time 11.3919 seconds
(without alignments)
1914.721 Million cell updates/sec

Title: US-09-896-301-7
Perfect score: 1275
Sequence: 1 DYCGMQSHATFYGGGDSAG.....AYNLVPSNMWQFGQTEGPQF 227

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: PIR_71:*
2: PIR1:*
3: PIR2:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1265	99.2	250	2	T10079
2	1092	85.6	232	2	T09821
3	1087	85.3	232	2	T09818
4	1085	85.1	232	2	T09826
5	1077	84.5	232	2	T09825
6	1072.5	84.1	253	2	T84831
7	1041	81.6	246	2	T04175
8	1041	81.6	251	2	T03298
9	1041	81.6	258	2	T09786
10	1027	80.5	237	2	T50654
11	1022.5	80.2	255	2	T50656
12	1019	79.9	248	2	C84444
13	985	77.3	257	2	D84820
14	960	75.3	260	2	T47689
15	954	74.8	262	2	T02530
16	947	74.3	261	2	T07630
17	945	74.1	257	2	T50658
18	940	73.7	255	2	T06573
19	939	73.6	257	2	T02727
20	939	73.6	259	2	T50653
21	937	73.5	258	2	S53082
22	935	73.3	264	2	T50659
23	934	73.3	258	2	T48247
24	934	73.3	260	2	T08016
25	924	72.5	262	2	T50660
26	889.5	69.8	255	2	T50655
27	874	68.5	258	2	T10083
28	859.5	67.4	261	2	T03737
29	853.5	67.1	252	2	F86335

30	832	65.3	255	2	T02010
31	819.5	64.3	255	2	T03299
32	730	57.3	160	2	T09871
33	678	53.2	257	2	G36654
34	668	52.4	257	2	F86259
35	604	47.4	256	2	T05648
36	322.5	25.3	102	2	T09828
37	317	24.9	77	2	T09815
38	280	22.0	81	2	T09830
39	278	21.8	259	2	T50657
40	278	21.8	271	2	H84592
41	276	21.6	277	2	S48032
42	271	21.3	261	2	T04301
43	270.5	21.2	491	2	F96681
44	266.5	20.9	259	2	F84886
45	242	19.0	262	2	S38620

ALIGNMENTS

RESULT 1

T10079
expansin S1 precursor - cucumber
C:Species: Cucumis sativus (cucumber)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000

R:Shcherban, T.Y.; Shi, J.; Durachko, D.M.; Gullitman, M.J.; McQueen-Mason, S.J.; Shi
Proc. Natl. Acad. Sci. U.S.A. 92, 9245-9249, 1995

A:Title: Molecular cloning and sequence analysis of expansins--a highly conserved, mu
A:Reference number: 214894; MUID:96016146

A:Accession: T10079
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-250 <SHC>

A:Cross-references: EMBL:U30382; NID:q1040874; PIDN:AB37746.1; PID:q1040875
A:Experimental source: cultivar Burpee Pickler
A:Genetics:

C:Function: A:Description: mediates cell wall extension
A:Superfamily: expansin
F:1-23/Domain: signal sequence #status predicted <SIG>

F:24-250/Product: expansin #status predicted <MAT>

Query Match 99.2%: Score 1265; DB 2; Length 250;
Best Local Similarity 99.1%: Pred. No. 3.9e-102;
Matches 225; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY	1	DYCGMQSHATFYGGGDSAGTGMGACGYGNYSGCYGTFVTALSTALPNNNGSGCACFEM	60
DB	24	DYCGMQSHATFYGGGDSAGTGMGACGYGNYSGCYGTFVTALSTALPNNNGSGCACFEM	83
QY	61	TCTNDPKMCLPCTIRYATNFCPPNFALPNDGNCNPLQHFDAEPAFLQIAYRAGI	120
DB	84	TCTNDPKMCLPCTIRYATNFCPPNFALPNNNGCNPPLQHFDAEPAFLQIAYRAGI	143
QY	121	VVVSRRRVCMKKGVRFTINSHFNVLVTINVGAGADYHVSITKSGRTGQMSRWMC	180
DB	144	VVVSRRRVCMKKGVRFTINSHFNVLVTINVGAGADYHVSITKSGRTGQMSRWMC	204
QY	181	QNMOSNNYLNQGLSFQVTLSDGRITLTAIVNLPVSNMFGQTEGPQF	227
DB	204	QNMOSNNYLNQGLSFQVTLSDGRITLTAIVNLPVSNMFGQTEGPQF	250

RESULT 2

T09821
expansin (clone pTexp3) - loblolly pine (fragment)

C:Species: Pinus taeda (loblolly pine)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 26-May-2000
C:Accession: T09821
R:Hutchinson, K.W.; Slinger, P.B.; Diaz-Sala, C.; Greenwood, M.S.

Sun Oct 13 14:19:16 2002

us-09-896-301-7.rpr

Page 6

```

Db      214 QNWQSNVILIGQSLSFRTASDRRSTSMWVAPATWQFGQTFSGKNE 260
          ||||| | ||||| | : ||| : : ||||| : |

```

Search completed: October 11, 2002, 15:01:01
Job time : 11.3919 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 11, 2002, 14:54:05 ; Search time 8.87897 Seconds
(without alignments)
624.466 Million cell updates/sec

Title: US-09-896-301-7

Perfect score: 1275

Sequence: 1 DYGMQSGHATFYGGGADSG.....AYNLVPSNMQFGQTEGPQF 227

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: Issued_Patents_AA:*
2: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1275	100.0	227	4	US-09-092-160-7
2	1021.5	80.1	226	2	US-08-440-517A-6
3	1021.5	80.1	226	4	US-09-092-160-6
4	1021	80.1	228	2	US-08-440-517A-2
5	1021	80.1	228	4	US-09-092-160-2
6	962	75.5	225	2	US-08-440-517A-5
7	962	75.5	225	4	US-09-092-160-5
8	857.5	67.3	222	2	US-08-440-517A-3
9	857.5	67.3	222	4	US-09-092-160-3
10	839	65.6	225	4	US-09-362-642-2
11	837	65.6	225	2	US-08-845-539-2
12	797	62.5	179	2	US-08-845-539-6
13	797	62.5	179	4	US-09-362-642-6
14	767.5	60.2	227	2	US-08-440-517A-4
15	767.5	60.2	227	4	US-09-092-160-4
16	732	57.4	167	2	US-09-362-642-4
17	732	57.4	167	4	US-08-845-539-4
18	245.5	19.3	261	1	US-07-971-096-2
19	245.5	19.3	261	1	US-08-175-096-2
20	233	18.3	246	4	US-08-441-507-21
21	231.5	18.2	263	1	US-07-971-096-4
22	231.5	18.2	263	1	US-08-175-096-4
23	231.5	18.2	263	4	US-08-413-974-6
24	231.5	18.2	263	4	US-08-434-418-6
25	231.5	18.2	263	4	US-08-433-288-6
26	231.5	18.2	263	4	US-08-174-739A-6
27	231	18.1	272	4	US-08-441-507-15

28	204	16.0	245	4	US-08-441-507-24	Sequence 24, Appl
29	173	13.6	197	4	US-08-441-507-5	Sequence 5, Appl
30	173	13.6	200	4	US-08-441-507-4	Sequence 4, Appl
31	139.5	10.9	122	4	US-08-441-507-23	Sequence 23, Appl
32	119.5	9.4	145	4	US-08-413-974-4	Sequence 4, Appl
33	119.5	9.4	145	4	US-08-434-418-4	Sequence 4, Appl
34	119.5	9.4	145	4	US-08-433-288-4	Sequence 4, Appl
35	119.5	9.4	145	4	US-08-174-739A-4	Sequence 4, Appl
36	100.5	7.9	54	3	US-08-750-419A-12	Sequence 12, Appl
37	95.5	7.5	145	4	US-08-441-507-16	Sequence 16, Appl
38	93.5	7.3	145	4	US-08-441-507-2	Sequence 2, Appl
39	86.5	6.8	138	4	US-08-441-507-7	Sequence 7, Appl
40	84.5	6.6	633	3	US-09-041-991A-8	Sequence 8, Appl
41	84.5	6.6	633	3	US-09-041-991A-10	Sequence 10, Appl
42	81.5	6.4	402	2	US-08-491-988-9	Sequence 9, Appl
43	79	6.2	402	1	US-08-460-806-11	Sequence 11, Appl
44	79	6.2	402	1	US-08-325-630-11	Sequence 11, Appl
45	78.5	6.2	418	4	US-09-254-733-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1
US-09-092-160-7
Sequence 7, Application US/09092160C
Patent No. 6255466
GENERAL INFORMATION:
APPLICANT: Cosgrove, Daniel J
APPLICANT: McQueen-Mason, Simon
APPLICANT: Guiltinan, Mark J
APPLICANT: Shcherban, Tatyana
APPLICANT: Shi, Jun
TITLE OF INVENTION: PURIFIED EXPANSIN PROTEINS
FILE REFERENCE: 1194/1C114053
CURRENT APPLICATION NUMBER: US/09/092,160C
EARLIER FILING DATE: 1998-06-05
EARLIER APPLICATION NUMBER: 08/440,517
EARLIER FILING DATE: 1995-05-12
EARLIER APPLICATION NUMBER: 08/242,090
EARLIER FILING DATE: 1994-05-12
EARLIER APPLICATION NUMBER: 08/060,944
EARLIER FILING DATE: 1993-05-12
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 7
LENGTH: 227
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: cucumber
US-09-092-160-7
Query Match 100.0%; Score 1275; DB 4; Length 227;
Best Local Similarity 100.0%; Pred. No. 2e-11b;
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DYGMQSGHATFYGGGADSGATMGACGYNLVLTNNVGAAGDVSVSIKSGRTGWSNRNWG 60
Db 1 DYGMQSGHATFYGGGADSGATMGACGYNLVLTNNVGAAGDVSVSIKSGRTGWSNRNWG 60
QY 1 TCTNDPKKCLGTTTAVTATNCPNFALPNDGCMCPDLPDHPMAEPALQIOTVYRAGT 120
Db 1 TCTNDPKKCLGTTTAVTATNCPNFALPNDGCMCPDLPDHPMAEPALQIOTVYRAGT 120
QY 121 VVPSRRVPCMKKGGVRETTNGHSYFNLVLTNNVGAAGDVSVSIKSGRTGWSNRNWG 180
Db 121 VVPSRRVPCMKKGGVRETTNGHSYFNLVLTNNVGAAGDVSVSIKSGRTGWSNRNWG 180
QY 181 QNMOSNNLNGOGLSFOVTLSDGRTLTAYNLVPSNMQFGQTEGPQF 227
Db 181 QNMOSNNLNGOGLSFOVTLSDGRTLTAYNLVPSNMQFGQTEGPQF 227

RESULT 2

US-08-440-517A-6

Sequence 6, Application US/08440517A

Patent No. 5959082

GENERAL INFORMATION:

APPLICANT: COSGROVE, DANIEL J.;

APPLICANT: GULLITMAN, MARK;

APPLICANT: SHCHERBAN, TATYANA;

APPLICANT: SHI, JUN

TITLE OF INVENTION: PURIFIED EXPANSIN PROTEINS

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: INTELLECTUAL PROPERTY OFFICE, THE

STREET: 113 TECHNOLOGY CENTER

CITY: UNIVERSITY PARK

STATE: PENNSYLVANIA

COUNTRY: UNITED STATES OF AMERICA

ZIP: 16802-7000

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK

COMPUTER: NEC 286

OPERATING SYSTEM: DOS

SOFTWARE: WORDPERFECT 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/440,517A

FILING DATE:

CLASSIFICATION: 530

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 226

TYPE: AMINO ACID

TOPOLOGY: UNKNOWN

US-08-440-517A-6

US-08-440-517A-6

Query Match

Best Local Similarity 80.1%; Score 1021.5; DB 2; Length 226;

Matches 177; Conservative 21; Mismatches 28; Indels 1; Gaps 1;

Query 1 DYGMOSGHAATFYGGDASGTMGACGYGNYLSOGYGTNTVALSTALFNNGISGACFEM 60

Db 1 DYSWOSAHATFYGGDASGTMGTCGYGNLYSTGY-TNTAALSTVLFNDGACRSCYEL 59

QY 61 TCTNDPKMCLPCTIRVATATNCPNFALPNDGWCNPLQHFDMAPFAILOIAOYRAGI 120

Db 60 RCDNDGWCMLPGSVTYTATNLCPNVALPNDGWCNPPRPHFDMAPFAILOIGYRAGI 119

QY 121 VPVSFRVPCMKRGVRFITNGHSYFNVLVTNVGAGDVHSYKSGRTGQMSRNMW 180

Db 120 VPVSFRVPCMKRGVRFITNGHSYFNVLVTNVGAGDVHSYKSGRTGQMSRNMW 179

QY 181 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 227

Db 180 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 226

QY 181 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 227

Db 180 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 226

QY 181 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 227

Db 180 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 226

QY 181 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 227

Db 180 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 226

QY 181 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 227

Db 180 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 226

QY 181 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 227

Db 180 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 226

QY 181 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 227

Db 180 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 226

QY 181 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 227

Db 180 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 226

QY 181 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 227

Db 180 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 226

QY 181 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 227

Db 180 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 226

QY 181 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 227

Db 180 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 226

QY 181 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 227

Db 180 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 226

QY 181 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 227

Db 180 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 226

QY 181 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 227

Db 180 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 226

QY 181 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 227

Db 180 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 226

QY 181 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 227

Db 180 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 226

QY 181 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 227

Db 180 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 226

EARLIER FILING DATE: 1995-05-12

EARLIER APPLICATION NUMBER: 08/242,090

EARLIER FILING DATE: 1994-05-12

EARLIER APPLICATION NUMBER: 08/060,944

EARLIER FILING DATE: 1993-05-12

NUMBER OF SEQ ID NOS: 7

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 6

LENGTH: 226

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Arabidopsis

OTHER INFORMATION: expansin

US-09-092-160-6

US-09-092-160-6

Query Match

Best Local Similarity 80.1%; Score 1021.5; DB 4; Length 226;

Matches 177; Conservative 21; Mismatches 28; Indels 1; Gaps 1;

Query 1 DYGMOSGHAATFYGGDASGTMGACGYGNYLSOGYGTNTVALSTALFNNGISGACFEM 60

Db 1 DYSWOSAHATFYGGDASGTMGTCGYGNLYSTGY-TNTAALSTVLFNDGACRSCYEL 59

QY 61 TCTNDPKMCLPCTIRVATATNCPNFALPNDGWCNPLQHFDMAPFAILOIAOYRAGI 120

Db 60 RCDNDGWCMLPGSVTYTATNLCPNVALPNDGWCNPPRPHFDMAPFAILOIGYRAGI 119

QY 121 VPVSFRVPCMKRGVRFITNGHSYFNVLVTNVGAGDVHSYKSGRTGQMSRNMW 180

Db 120 VPVSFRVPCMKRGVRFITNGHSYFNVLVTNVGAGDVHSYKSGRTGQMSRNMW 179

QY 181 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 227

Db 180 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 226

QY 181 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 227

Db 180 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 226

QY 181 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 227

Db 180 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 226

QY 181 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 227

Db 180 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 226

QY 181 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 227

Db 180 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 226

QY 181 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 227

Db 180 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 226

QY 181 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 227

Db 180 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 226

QY 181 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 227

Db 180 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 226

QY 181 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 227

Db 180 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 226

QY 181 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 227

Db 180 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 226

QY 181 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 227

Db 180 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 226

QY 181 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 227

Db 180 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 226

QY 181 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 227

Db 180 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 226

QY 181 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 227

Db 180 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 226

QY 181 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 227

Db 180 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 226

QY 181 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 227

Db 180 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 226

QY 181 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 227

Db 180 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 226

QY 181 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 227

Db 180 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 226

QY 181 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 227

Db 180 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 226

QY 181 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 227

Db 180 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 226

QY 181 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 227

Db 180 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 226

QY 181 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 227

Db 180 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 226

QY 181 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 227

Db 180 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 226

QY 181 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 227

Db 180 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 226

QY 181 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 227

Db 180 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 226

QY 181 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 227

Db 180 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 226

QY 181 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 227

Db 180 QNMOSNNTLNGGSLSFQVAVSDGRTVTSNNVPAWGQFQTEGGQF 226

Best Local Similarity 78.8%; Pred. No. 2,6e-93;
Matches 175; Conservative 23; Mismatches 24; Indels 0; Gaps 0;

Qy	3	GCWDSGCHATFTYGGCGDASGTMGACGCGNLY	SSQSGTNTYALSTALPNNLNSGCAECMEYC	62
Db	3	GGWNAHATATYGGGDSASGTGAGACGCGNLY	SSQSGTNTYALSTALPNNLNSGCAECMEYC	62
Qy	63	TNDPKMCLPCTIRATATATNFCPCPFALP	PNDDGWCNCNPLQHFDMAEPALQAOYRAGIYP	122
Db	63	QNDCKMCLPSTIVYATATNFCPCPNPALP	NNAGWCNCNPPQHFHDLSSQVYFORINQYRAGIYP	122
Qy	123	VSEFRVPCMKKGGVRFPTINGSYFNLVLT	NVAGGADVHSYLSIKGRTGQMSKRMNGON	182
Db	123	VAYRVPVCVRGGCIRFTINGSYFNLVLT	NVAGGADVHSAMKGSRTGQMSKRMNGON	182
Qy	183	WQSNNTYLANGGLSPQVLYLSGCRLLTAV	NLVPSMWQGGQYREG	224
Db	183	WQSNSTYLANGGLSKFTVTSQGITIVSN	XNAQMSFGQFTTG	224

RESULT 5
US-09-092-160-2

Sequence 2, Application US/09092160C
Patent No. 6255466
GENERAL INFORMATION:
APPLICANT: Cosgrove, Daniel J
APPLICANT: McQueen-Mason, Simon
APPLICANT: Gullinan, Mark J
APPLICANT: Shcherban, Tatyana
APPLICANT: Shi, Jun
TITLE OF INVENTION: PURIFIED EXPANSIN PROTEINS
FILE REFERENCE: 1194/1011US3
CURRENT APPLICATION NUMBER: US/09/092,160C
CURRENT FILING DATE: 1998-06-05
EARLIER APPLICATION NUMBER: 08/440,517
EARLIER FILING DATE: 1995-05-12
EARLIER APPLICATION NUMBER: 08/242,090
EARLIER FILING DATE: 1994-05-12
EARLIER APPLICATION NUMBER: 08/060,944
EARLIER FILING DATE: 1993-05-12
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 228
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: rice expansin
FEATURE:
NAME/KEY: UNSURE
LOCATION: 211
OTHER INFORMATION: Xaa is unknown or other.
US-09-092-160-2

Query Match	80.1%;	Score 1021;	DB 4;	Length 228;
Best Local Similarity	78.8%;	Pred. No. 2.6e-93;		
Matches 175;	Conservative 23;	Mismatches 24;	Indels 0;	Gaps 0;

[illegible]

RESULT 6
US-08-440-517A-5

Sequence 5, Application US/08440517A
Patent No. 5959082

GENERAL INFORMATION:

APPLICANT: COSGROVE, DANIEL J.,
APPLICANT: GULLITMAN, MARK,
APPLICANT: SCHERMBAN, TATIANA,
APPLICANT: SHI, JUN

TITLE OF INVENTION: PURIFIED EXPANSIN PROTEINS

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: INTELLECTUAL PROPERTY OFFICE, THE
ADDRESSEE: PENNSYLVANIA STATE UNIVERSITY
STREET: 113 TECHNOLOGY CENTER
CITY: UNIVERSITY PARK
STATE: PENNSYLVANIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 16802-7000

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK

COMPUTER: NEC 286

OPERATING SYSTEM: DOS

SOFTWARE: WORDPERECT 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/440,517A

FILING DATE:

CLASSIFICATION: 530

INFORMATION FOR SEQ. ID NO.: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 225

TYPE: AMINO ACID

TOPOLOGY: UNKNOWN

US-08-440-517A-5

Query Match	75.5%	Score 962;	DB 2;	Length 11	225;
Best Local Similarity	75.0%;	Pred. No. 1.7e-87;			
Matches 171;	Conservative 26;	Mismatches 27;	Indels 4;	Gaps 2;	

```

QY      1 DYGMQSHATFFGGGADSGTMGACGYNLYSGYGTNTVALSTALFNNGLSCGACFEM 60
        1 1111111111111111111111111111111111111111111111111111111
Db      1 DNGGERRHATFFGGADSGTMGACGYNLHSGYGLQTLAALSTALFNSGGQCGACFEL 60
QY      61 TCTNDPKMKCLGTRVYATNTCPENFALPDDGWCMPPLQHFEMAPFLQIAQVRAGI 120
        1111111111111111111111111111111111111111111111111111111
Db      61 TCEDDEPENCIGSTIIV---RYNLNANFALANDNGWCMPPLKHFEDLAPFLQIAQVRAGI 117
QY      121 VVFSRRRPPCKMKKGVRPTINGHSYFNLVLTNTVGAGADHSYVSIKSGRTG--WQSMRRW 179
        1111111111111111111111111111111111111111111111111111111
Db      118 VVFAARRRPPCKEKGIRRTINGNRYFDELVLTNTVGAGADIRAVSLKSKTDQWQSMRRW 177
QY      180 GGNWQSNNTLNGQGLSFQVLTLSDCRRTLTAIYLVPSNMQFGQTYEGPQF 227
        1111111111111111111111111111111111111111111111111111111
Db      178 GGNWQSNNTLNGQGLSFQVLTLSDCRRTIVSYXVYVHDMQFGQTEGCGPF 225

```

```

RESULT 7
US-09-092-160-5
; Sequence 5, Application US/09092160C
; Patent No. 6255466
; GENERAL INFORMATION:
; APPLICANT: Cosgrove, Daniel J
; APPLICANT: McQueen-Mason, Simon
; APPLICANT: Gullinan, Mark J
; APPLICANT: Shcherban, Tatyana
; APPLICANT: Shi, Jun
; TITLE OF INVENTION: PURIFIED EXPANSIN PROTEINS
; FILE REFERENCE: 1194/1C114US3
; CURRENT APPLICATION NUMBER: US/09/092.160C
; CURRENT FILING DATE: 1998-06-05
; EARLIER APPLICATION NUMBER: 08/440,517

```

```

: RESULT 7
: US-09-092-160-5
: Sequence 5, Application US/09092160C
: Patent NO. 6255466
: GENERAL INFORMATION:
: APPLICANT: Cosgrove, Daniel J
: APPLICANT: McQueen-Mason, Simon
: APPLICANT: Gullinan, Mark J
: APPLICANT: Shcherban, Tatyana
: APPLICANT: Shi, Jun
: TITLE OF INVENTION: PURIFIED EXPANSIN PROTEINS
: FILE REFERENCE: 1194/1C14US3
: CURRENT FILING DATE: 1998-06-05
: CURRENT APPLICATION NUMBER: US/09/092,160C
: EARLIER APPLICATION NUMBER: 08/440,517

```



```

RESULT 10
US-09-362-642-2
; Sequence 2, Application US/09362642
; Patent No. 6350935
; GENERAL INFORMATION:
; APPLICANT: Bennett, Alan B.
; APPLICANT: Rose, Jocelyn K.C.
; TITLE OF INVENTION: The Regents of the University of California
; TITLE OF INVENTION: Fruit-Specific and Ripening Regulation Expansin Genes
; FILE REFERENCE: 023070-078210US
; CURRENT APPLICATION NUMBER: US/09/362,642
; CURRENT FILING DATE: 1999-07-27
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Lycopersicon esculentum cv. T5

Query Match          65.8%; Score 839; DB 4; Length 225;
Best Local Similarity 74.7%; Pred. No. 2,4e-75;
Matches 145; Conservative 21; Mismatches 24; Indels 4; Gaps 2;

QY 3 GGMOSGHAIFYGGDASCTMGACGYGNLYSQGYGTNTALSTALFNNGLSGACFEWTC 62
DB 31 GSWETAHATFYGSDASCTMGACGYGNLYSQGYGTNTALSTALFNNGLSGACFEWTC 62
QY 63 TNDP--KWCLPG--TIVTATNFCPPNFPALPNDGCMCNPPLOHFDMAEPALQIAQYRA 118
DB 91 TNPMMKCLPGNPSILITATNFCPPNFPALPNDGCMCNPPRPHFLDAMPFLKLAQYRA 150
QY 119 GIVPVSFRVPCMKKGVRFTNGHSYENLVLTITNAGAGDVHSVTSKGRSTGQMSNRN 178
DB 151 GIVPVTTRIRICRKGGRFTNGHSYENLVLTITNAGAGDVHSVTSKGRSTGQMSNRN 178
QY 179 WGMQMSNMYLNGQ 192
DB 211 WGMQMSNAVLNGQ 224

RESULT 11
US-08-845-539-2
; Sequence 2, Application US/08845539
; Patent No. 5929303
; GENERAL INFORMATION:
; APPLICANT: Bennett, Alan B.
; APPLICANT: Rose, Jocelyn K.C.
; TITLE OF INVENTION: Fruit-Specific and Ripening-Regulation
; TITLE OF INVENTION: Expansin Genes to Control Fruit Texture and Softening
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/845,539
; FILING DATE: 25-APR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774

```

```

; REFERENCE/DOCKET NUMBER: 023070-078200US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 225 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-845-539-2

Query Match          65.6%; Score 837; DB 2; Length 225;
Best Local Similarity 74.7%; Pred. No. 3.9e-75;
Matches 145; Conservative 20; Mismatches 25; Indels 4; Gaps 2;

QY 3 GGMOSGHAIFYGGDASCTMGACGYGNLYSQGYGTNTALSTALFNNGLSGACFEWTC 62
DB 31 GSWETAHATFYGSDASCTMGACGYGNLYSQGYGTNTALSTALFNNGLSGACFEWTC 62
QY 63 TNDP--KWCLPG--TIVTATNFCPPNFPALPNDGCMCNPPLOHFDMAEPALQIAQYRA 118
DB 91 TNPMMKCLPGNPSILITATNFCPPNFPALPNDGCMCNPPRPHFLDAMPFLKLAQYRA 150
QY 119 GIVPVSFRVPCMKKGVRFTNGHSYENLVLTITNAGAGDVHSVTSKGRSTGQMSNRN 178
DB 151 GIVPVTTRIRICRKGGRFTNGHSYENLVLTITNAGAGDVHSVTSKGRSTGQMSNRN 178
QY 179 WGMQMSNMYLNGQ 192
DB 211 WGMQMSNAVLNGQ 224

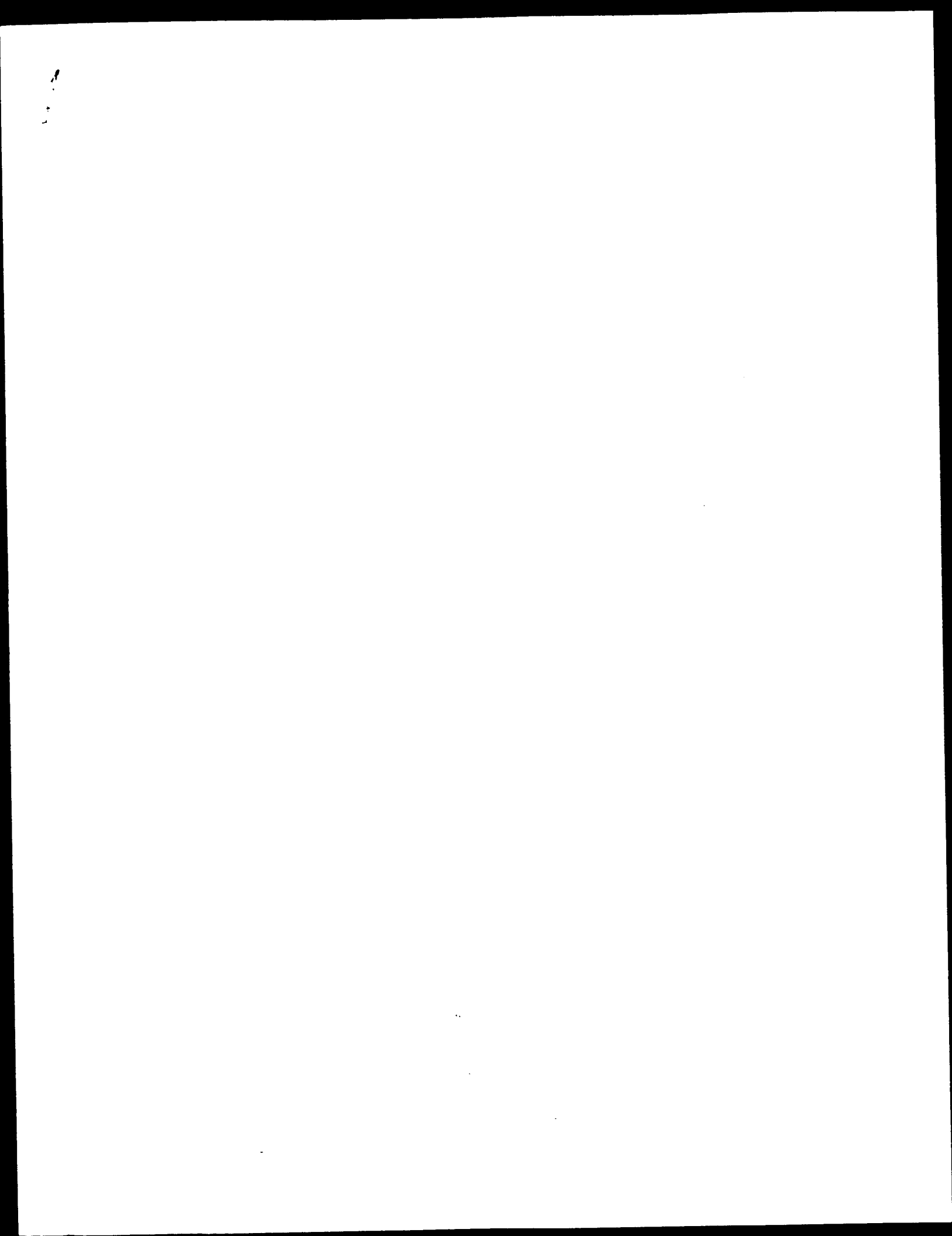
RESULT 12
US-08-845-539-6
; Sequence 6, Application US/08845539
; Patent No. 5929303
; GENERAL INFORMATION:
; APPLICANT: Bennett, Alan B.
; APPLICANT: Rose, Jocelyn K.C.
; TITLE OF INVENTION: Fruit-Specific and Ripening-Regulation
; TITLE OF INVENTION: Expansin Genes to Control Fruit Texture and Softening
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/845,539
; FILING DATE: 25-APR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-078200US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 179 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-845-539-6

```

US-09-092-160-4

Query Match 60.2%: Score 767.5; DB 4; Length 227;
Best Local Similarity 64.0%: Pred No. 2.9e-68;
Matches 137; Conservative 27; Mismatches 47; Indels 3; Gaps 2;
QY 6 QSGHATFYGGDASGTMGAGCYGNLYSGYGTNTVALSTALFNNGLSGACFEWTC-TN 64
DB 6 QSAFAFTEYGGKDSCGTMGAGCYGNLYNAGYGLYNALSSALFNNGAMGACYTITCDTS 65
QY 65 DPKWCLPG--TTRVTATNFCPPNFALPNDGGMCPPLQHFDMAPFLQTAQYRAGIYP 122
DB 66 QTRKCKPFGNSITITATNLCPNMLPSNSGMCNPPLXHPDMSQPAWENIAVYQAGIYP 125
QY 123 VSEFRVPCMKKGVRFTINGHSYFNLVLTNVGAGADYHSYIKSRTGMQMSRNMGN 182
DB 126 VNYKRVXPORSRGIRFATSGHDYFELYVTNVGSGCYVAQMSIKSNTGMMSRNMGN 185
QY 183 WQSNMYLNGGSLFOYVTLSDGRTLTAYNLVPSPNW 216
DB 186 WQSNMYLNGGSLFTVQLDDGRRKYTAMNXPXNW 219

Search completed: October 11, 2002, 15:02:08
Job time : 9.87897 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 11, 2002, 14:49:55 : Search time 25.1292 Seconds
(without alignments)
1003.367 Million cell updates/sec

Title: US-09-896-301-7

Perfect score: 1275

Sequence: 1 DYGMOSGHATFYCGDASG.....AYNLVPSNMQFGQTRBPQF 227

Scoring table:

BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: A_Geneseq_032802.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1265	99.2	227	17	AR94527
2	1072.5	84.1	253	21	AA625443
3	1072.5	84.1	253	21	AA646483
4	1031	80.9	241	21	AA606546
5	1031	80.9	241	21	AA651633
6	1031	80.9	241	21	AA651647
7	1031	80.9	249	21	AA606545
8	1031	80.9	249	21	AA651632
9	1031	80.9	249	21	AA651646
10	1031	80.9	259	21	AA606544
11	1031	80.9	259	21	AA651631

12	1031	80.9	280	21	AA651645	Arabidopsis thalia
13	1021	80.1	228	17	AA94528	Rice expansin. Or
14	1019	79.9	251	21	AA623852	Arabidopsis thalia
15	1019	79.9	251	21	AA643343	Arabidopsis thalia
16	1019	79.9	253	21	AA623851	Arabidopsis thalia
17	1019	79.9	253	21	AA643342	Arabidopsis thalia
18	1019	79.9	281	21	AA643341	Arabidopsis thalia
19	1019	79.9	282	21	AA623850	Arabidopsis thalia
20	1012.5	79.4	227	17	AA94532	Arabidopsis thalia
21	994	78.0	250	22	AA600414	Arabidopsis thalia
22	985	77.3	257	21	AA636445	Tomato seed expans
23	970.5	76.1	207	21	AA625444	Arabidopsis thalia
24	970.5	76.1	207	21	AA646484	Arabidopsis thalia
25	965	75.7	249	21	AA605453	Arabidopsis thalia
26	965	75.7	249	21	AA605452	Arabidopsis thalia
27	965	75.7	255	21	AA605451	Arabidopsis thalia
28	962	75.5	225	17	AA94531	Arabidopsis thalia
29	960	75.3	242	21	AA636570	Arabidopsis thalia
30	960	75.3	249	21	AA636569	Arabidopsis thalia
31	960	75.3	253	21	AA636568	Arabidopsis thalia
32	954	74.8	252	21	AA629931	Arabidopsis thalia
33	954	74.8	273	21	AA629930	Arabidopsis thalia
34	939	73.6	255	21	AA630325	Arabidopsis thalia
35	939	73.6	257	21	AA630324	Arabidopsis thalia
36	938	73.6	253	22	AA600412	Arabidopsis thalia
37	934	73.3	250	21	AA609622	Tomato seed expans
38	934	73.3	258	21	AA609621	Arabidopsis thalia
39	934	73.3	280	21	AA609620	Arabidopsis thalia
40	895	70.2	210	21	AA636446	Arabidopsis thalia
41	889.5	69.8	221	21	AA615694	Arabidopsis thalia
42	889.5	69.8	225	21	AA651013	Arabidopsis thalia
43	889.5	69.8	251	21	AA615693	Arabidopsis thalia
44	889.5	69.8	255	21	AA651012	Arabidopsis thalia
45	866	67.9	210	21	AA629932	Arabidopsis thalia

ALIGNMENTS

RESULT 1
AA94527
ID AA94527 standard; Protein; 227 AA.
XX
AC AA94527;
XX
DT 08-JUL-1996 (first entry)
XX
DE Cucurbit expansin-29.
XX
KW Cucurbit expansin-29; plant cell wall; cellulose; paper recycling; de-linking;
XX
KW polysaccharide; cucumber.
XX
OS Cucurbit sativus var. Burpee Pickler.
XX
PN AU9540262-A.
XX
PD 04-APR-1996.
XX
PF 12-MAY-1994; 94AU-0068320.
XX
PR 12-MAY-1995; 95US-0440517.
XX
PR 12-MAY-1993; 93US-0060944.
XX
PA (PENN-) PENN STATE RRS FOUND.
XX
PI Cosgrove DJ, McQueen-Mason S;
XX
DR WPI, 1996-201150/21.
XX
PT N-PSDB; AAT13320.
XX
PT Expansin proteins which alter the mechanical strength of
XX poly:saccharide(s) - useful in paper mfr. and recycling

PS Claim 7, page 30; 60pp; English.
XX
CC Cucurbit expansin-29 (AA94527) is a member of a novel class of
CC proteins that catalyse the extension of plant cell walls and the
CC weakening of the hydrogen bonds in pure cellulose. It can be obtd.
CC by expression of an isolated cDNA clone (see AAT13320) in bacterial or
CC other host cells. Expansin proteins have also been identified in oat
CC coleoptiles, in Arabidopsis (see AA94530-32) and in rice (AA94528-29),
CC and appear to be broadly distributed throughout the plant kingdom.
CC Expansins can be used e.g. in the mfr. de-linking and recycling of
CC paper, in the textile industry, to aid delignification processes, to
CC alter gel mechanical strength, etc.
XX
SQ Sequence 227 AA:
Query Match 99.2%; Score 1265; DB 17; Length 227;
Best Local Similarity 99.1%; Pred. No. 1,1e-119;
Matches 225; Conservative 2; Mismatches 0; Indels 0; Gaps 0:
QY 1 DYGMQSGHATFYGGDASCTMGAGCGYGNLYSGYGTNTVALSTALFNNGISCGACFEM 60
1 DYGMQSGHATFYGGDASCTMGAGCGYGNLYSGYGTNTVALSTALFNNGISCGACFEM 60
DB 61 TCTNDRKWCIPGTRITATNFCPPNLPALPDGSGCNPLDHPDMEPAFLQIAOYRACI 120
61 TCTNDRKWCIPGTRITATNFCPPNLPALPDGSGCNPLDHPDMEPAFLQIAOYRACI 120
DB 61 TCTNDRKWCIPGTRITATNFCPPNLPALPDGSGCNPLDHPDMEPAFLQIAOYRACI 120
61 TCTNDRKWCIPGTRITATNFCPPNLPALPDGSGCNPLDHPDMEPAFLQIAOYRACI 120
QY 121 VPSFRRVPCMKKGAVFTINGHSYENLVLTINVGAGADVHSYISIKSRTGWSNRMG 180
121 VPSFRRVPCMKKGAVFTINGHSYENLVLTINVGAGADVHSYISIKSRTGWSNRMG 180
DB 121 VPSFRRVPCMKKGAVFTINGHSYENLVLTINVGAGADVHSYISIKSRTGWSNRMG 180
121 VPSFRRVPCMKKGAVFTINGHSYENLVLTINVGAGADVHSYISIKSRTGWSNRMG 180
QY 181 QNMOSNNYLVNGGILSFQVTLSDGRTLTATLVPSNMQFGQTEGPOE 227
181 QNMOSNNYLVNGGILSFQVTLSDGRTLTATLVPSNMQFGQTEGPOE 227
DB 181 QNMOSNNYLVNGGILSFQVTLSDGRTLTATLVPSNMQFGQTEGPOE 227
181 QNMOSNNYLVNGGILSFQVTLSDGRTLTATLVPSNMQFGQTEGPOE 227
RESULT 2
AAG25443
ID AAG25443 standard; Protein: 253 AA.
AC AAG25443:
XX 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 29510.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 99US-0131449.
PR 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0138453.
PR 17-JUN-1999; 99US-0138492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0138459.
PR 18-JUN-1999; 99US-0138460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0138763.
PR 21-JUN-1999; 99US-0138817.
PR 22-JUN-1999; 99US-0138899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140921.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142877.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.

PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130892.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 14-MAY-1999; 99US-0134768.
PR 18-MAY-1999; 99US-0134941.
PR 19-MAY-1999; 99US-0135124.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140655.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142305.
PR 06-JUL-1999; 99US-0142380.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.

PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159684.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 25-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 26-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 84.1%; Score 1072.5; DB 21; Length 253;
Best Local Similarity 82.9%; Pred. No. 3.7e-100;
Matches 189; Conservative 17; Mismatches 21; Indels 1; Gaps 1;

OY 1 DYVGQSGHAFYGGGDSGTMGACGCGNLYSGCTGNTALSTALFNNGISGACPEM 60
DB 26 DDGGQGGHAFYGGGDSGTMGACGCGNLYSGCTGNTALSTALFNNGISGACPEM 85
OY 61 TCTNDPKMCLPPTJRTVATNFCPPNFALPNDGCGWCPNPLQHPMAEPAPFQIAOYRAGI 120
DB 86 KCNDPPKMCSTITVTATNFCPPNPGLSNDGCGWCPNPLQHPMAEPAPFQIAOYRAGI 145
OY 121 VPSFRARVPCKKGGVARTINGHSYFNLYLTNNGGACDVHSVSTKGSRT-GWQSMRNW 179
DB 146 VPSFRARVPCKKGGVARTINGHSYFNLYLTNNGGACDVHSVSTKGSRT-GWQSMRNW 205
OY 180 GNNQSNVYLVNGCGLSEFQVTLSDGRTLVAVNYLVPSNMQFGQTYGCPOR 227
DB 206 GNNQSNVYLVNGCGLSEFQVTLSDGRTLVAVNYLVPSNMQFGQTYGCPOR 253

RESULT 4
AAG06546
ID AAG06546 standard; Protein: 241 AA.

AC AAG06546;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 3357.

KM Protein identification: signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135124.
PR 24-MAY-1999; 99US-0135353.
PR 25-MAY-1999; 99US-0135629.
PR 27-MAY-1999; 99US-0136021.
PR 28-MAY-1999; 99US-0136392.
PR 03-JUN-1999; 99US-0136782.
PR 03-JUN-1999; 99US-0137222.
PR 04-JUN-1999; 99US-0137528.
PR 07-JUN-1999; 99US-0137502.
PR 08-JUN-1999; 99US-0137724.
PR 10-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 28-JUN-1999; 99US-0140695.
PR 29-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0140991.
PR 01-JUL-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0142977.
PR 14-JUL-1999; 99US-0143542.
PR 15-JUL-1999; 99US-0143624.
PR 16-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.

PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134258.
PR 14-MAY-1999; 99US-0134259.
PR 14-MAY-1999; 99US-0134261.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137528.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139493.
PR 18-JUN-1999; 99US-0139494.
PR 18-JUN-1999; 99US-0139495.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141847.
PR 01-JUL-1999; 99US-0142158.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.

PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144333.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151330.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154038.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.

XX	06-SEP-2000.	
PD		
XX	25-FEB-2000.	2000BP-0301439.
XX		
PR	25-FEB-1999;	9905-0121895.
PR	05-MAR-1999;	9905-0123180.
PR	09-MAR-1999;	9905-0125348.
PR	23-MAR-1999;	9905-0125788.
PR	25-MAR-1999;	9905-0126564.
PR	29-MAR-1999;	9905-0126785.
PR	01-APR-1999;	9905-0127652.
PR	06-APR-1999;	9905-0128234.
PR	08-APR-1999;	9905-0128714.
PR	16-APR-1999;	9905-0129945.
PR	19-APR-1999;	9905-0130077.
PR	21-APR-1999;	9905-0130449.
PR	23-APR-1999;	9905-0130510.
PR	23-APR-1999;	9905-0130891.
PR	28-APR-1999;	9905-0131449.
PR	30-APR-1999;	9905-0132048.
PR	30-APR-1999;	9905-0132407.
PR	04-MAY-1999;	9905-0132484.
PR	05-MAY-1999;	9905-0132485.
PR	06-MAY-1999;	9905-0132486.
PR	06-MAY-1999;	9905-0132487.
PR	07-MAY-1999;	9905-0132863.
PR	11-MAY-1999;	9905-0134256.
PR	14-MAY-1999;	9905-0134281.
PR	14-MAY-1999;	9905-0134291.
PR	14-MAY-1999;	9905-0134320.
PR	18-MAY-1999;	9905-0134768.
PR	19-MAY-1999;	9905-0134961.
PR	20-MAY-1999;	9905-0135124.

PR 21-MAY-1999; 99US-0135629.
PR 24-MAY-1999; 99US-0136021.
PR 25-MAY-1999; 99US-0136392.
PR 27-MAY-1999; 99US-0136392.

PR	01-MAR-1999	9905-0137222
PR	01-JUN-1999	9905-0137528
PR	03-JUN-1999	9905-0137502
PR	04-JUN-1999	9905-0137500
PR	07-JUN-1999	9905-0137724
PR	08-JUN-1999	9905-0138094
PR	10-JUN-1999	9905-0138540
PR	10-JUN-1999	9905-0138847
PR	10-JUN-1999	9905-0138319
PR	14-JUN-1999	9905-0138452
PR	16-JUN-1999	9905-0138453
PR	16-JUN-1999	9905-0138455
PR	17-JUN-1999	9905-0138492
PR	18-JUN-1999	9905-0138454
PR	18-JUN-1999	9905-0138455
PR	18-JUN-1999	9905-0138456
PR	18-JUN-1999	9905-0138457

PR	18-JUN-1999;	99US-0139460;
PR	18-JUN-1999;	99US-0139461;
PR	18-JUN-1999;	99US-0139462;
PR	18-JUN-1999;	99US-0139463;
PR	18-JUN-1999;	99US-0139750;
PR	18-JUN-1999;	99US-0139763;
PR	18-JUN-1999;	99US-0139817;
PR	22-JUN-1999;	99US-0139859;
PR	22-JUN-1999;	99US-0140353;
PR	23-JUN-1999;	99US-0140354;
PR	23-JUN-1999;	99US-0140695;
PR	24-JUN-1999;	99US-0140823;
PR	28-JUN-1999;	99US-0140991;
PR	29-JUN-1999;	99US-0141287;
PR	30-JUN-1999;	99US-0141287;
PR	01-JUL-1999;	99US-0141442;
PR	01-JUL-1999;	99US-0142154;

```

PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142927.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148365.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150566.
PR 27-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.

```

```

PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157153.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159337.
PR 18-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 21-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Query Match

Best Local Similarity 80.9%; Score 1031; DB 21; Length 241;

Matches 177; Conservative 25; Mismatches 23; Indels 0; Gaps 0;

```

QY 3 GCMQGHATFYGGDAGTCGACGAGNLYSGGTNTVALSTALFNNGISGACPEMTG 62
   ||| : |||||||
DB 16 GWTNAHATFYGGDAGTCGACGAGNLYSGGTNTVALSTALFNNGISGACPEMTG 75
   ||| : |||||||

QY 63 TNDPKMCLPGTIRYATNFPCPNFALPNDGCGMCPNLPDMPAPPAFLQIYRAGTVP 122
   ||| : |||||||
DB 76 ENDGKMCPLPGSIVVTATNFPCPNNALANNNGCMCPNLPDMPAPPAFLQIYRAGTVP 135
   ||| : |||||||

QY 123 VSPFRVPCMKKGGVFTINGHSYFNLYLVITNVGAGDVHSYKSRKQMSRNMCON 182
   ||| : |||||||
DB 136 VSTRVPCRRRGIRFTINGHSYFNLYLVITNVGAGDVHSYKSRVQMSRNMCON 195
   ||| : |||||||

QY 183 WQSNVYLMGGLSFQYTLSDGRTLVYNIYVPSNMWQRCQYRBPOR 227
   ||| : |||||||
DB 196 WQSNVYLMGGLSFQYTLSDGRTLVYNIYVPSNMWQRCQYRBPOR 240
   ||| : |||||||

```

RESULT 7

AAG06545 AAG06545 standard; Protein; 249 AA.

AC AAG06545;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment seq ID NO: 3356.

KW Protein identification: signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;

PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 23-OCT-1999; 99US-0160989.
PR 23-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161933.
PR 29-OCT-1999; 99US-0162142.

Query Match 80.9%; Score 1031; DB 21; Length 249;
Best Local Similarity 78.7%; Pred. No. 5, 7e-96;
Matches 177; Conservative 25; Mismatches 23; Indels 0; Gaps 0;

OY 3 GGMOSGHATFYGGDASGFMGACGYGNLYSQGYGTNTVALSTALFNNGLSGCAFEEMTC 62
DB 24 GCHINHAATFYGGDASGFMGACGYGNLYSQGYGTNTVALSTALFNNGLSGCAFEEMTC 83
OY 63 TNDPKWCLPGLTAVTATNFCPPNFALPNDGNCNPPLOHFDMAEPAFLIOAYRAGIYV 122
DB 84 ENDGKWCPLGSIYVATNFCPPNFALPNDGNCNPPLOHFDMAEPAFLIOAYRAGIYV 143
OY 123 VSFRRPCKAKKGVFTTNGHSYFNVLVTNNVGAGDVHVSFKSGSTQGSRRMNGON 162
DB 144 VSYRRVPCARRRGIRFTINGHSYFNVLVTNNVGAGDVHSAALIGSTVQAMSRMNGON 203
OY 183 WOSNNYLNCGSLFQYTLSDGRTLVAVNLVPSNMOPGQTEGPOF 227
DB 204 WOSNNYLNCGSLFQYTLSDGRTLVAVNLVPSNMOPGQTEGPOF 248

RESULT 8
AAGS1632
ID AAGS1632 standard; Protein; 249 AA.
AC AAGS1632;
XX
XX
DT 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 65550.
DE
XX Protein identification: signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000BP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135323.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.

PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140683.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147392.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.

PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 80.9%; Score 1031; DB 21; Length 249;
Best Local Similarity 78.7%; Pred. No. 5 7e-96;
Matches 177; Conservative 25; Mismatches 23; Indels 0; Gaps 0;

QY 3 GGNQSGHATFYGCGDASGTMGACGYNLYSQGYGTNTVALSTALFNNGLSGACFEWIC 62
DB 24 GGNINAHATFYGGDASGTMGACGYGNLYSQGYGTNTVALSTALFNNGLSGACFEIRC 83
QY 63 TNDPKCLPCTIRVTATNCPFPFALPNDGCGCNPLDHPMAEPALQIOTYRAGIYP 122
DB 84 ENDGKWCIPDSIVTATNCPFPFALPNDGCGCNPLDHPMAEPALQIOTYRAGIYP 143
QY 123 VSFRRVPCMKKGGVFTINGHSYFNVLITNNGAGDVHSYLSKSRGOSMRNMGON 182
DB 144 VSTRRVPCRRRGIRFTINGHSYFNVLITNNGAGDVHSYLSKSRIVWQMSRNMGON 203
QY 183 WQSNNTYLNQGLSFQYTLSDGRTITAYNLVPSNMDFGQTYEGRPOF 227
DB 204 WQSNNTYLNQGLSFQYTLSDGRTITAYNLVPSNMDFGQTYEGRPOF 248

RESULT 9
AAG51646

ID AAG51646 standard; Protein; 249 AA.
XX AAG51646;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 65569.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN Epi03405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 07-MAY-1999; 99US-0132486.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 07-JUN-1999; 99US-0137702.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138447.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139352.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0142977.
PR 15-JUL-1999; 99US-0143542.
PR 15-JUL-1999; 99US-0143624.
PR 16-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 12-AUG-1999; 99US-0148319.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.

```

PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155658.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 13-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159337.
PR 14-OCT-1999; 99US-0159363.
PR 18-OCT-1999; 99US-0159584.
PR 18-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161932.
PR 28-OCT-1999; 99US-0161933.
PR 29-OCT-1999; 99US-0162142.

```

Query Match 80.9%; Score 1031; DB 21; Length 249;
 Best Local Similarity 78.7%; Pred. No. 5,7e-96;
 Matches 177; Conservative 25; Mismatches 23; Indels 0; Gaps 0;

```

QY 3 GGNOSGATRYGGGADAGTGMGAGCGNLYSQGYGTVALSTALFNNGLSCGACFEPMC 62
DB 24 GGNATNATATYGGGADAGTGMGAGCGNLYSQGYGTVALSTALFNNGLSCGACFEIRC 83
QY 63 TNDKWCLEPGTIRVTATNFCPPNAPLNDGCGNPPLOHFDMAEPFLQIAYRAGIYP 122
DB 84 ENDKWCLEPGTIRVTATNFCPPNAPLNDGCGNPPLOHFDMAEPFLQIAYRAGIYP 143
QY 123 VSEFRVCMKKGVRFTINGSHYFNLYITNVGAGDVHSHVSIKSGRTGQMSRMNGON 182
DB 144 VSTRVFCRRKRGIRFTINGSHYFNLYITNVGAGDVHSHVSIKSGRTVQMSRMNGON 203
QY 183 WOSNNYINGGLSFQVTLSDGRTLTAVNLVPSNMQFGQTEGPOF 227

```

```

DB 204 WOSNNYINGGLSFQVTLSDGRTVSEFNAPAGMSYGGTFAGQF 248
RESULT 10
AACG06544
ID AACG06544 standard; Protein; 259 AA.
XX AACG06544;
AC
XX
XX 17-OCT-2000 (first entry)
DE
XX Arabidopsis thaliana protein fragment SEQ ID NO: 3355.
KW Protein identification: signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
OS Arabidopsis thaliana.
PN EPI033405-A2.
PD
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126284.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134370.
PR 14-MAY-1999; 99US-0134376.
PR 18-MAY-1999; 99US-0134941.
PR 19-MAY-1999; 99US-0135124.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 21-MAY-1999; 99US-0135629.
PR 24-MAY-1999; 99US-0136021.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 14-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.

```

PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139751.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140699.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 28-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 03-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 06-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 09-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 10-AUG-1999; 99US-0147935.
 PR 11-AUG-1999; 99US-0148171.
 PR 12-AUG-1999; 99US-0148319.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.

PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158222.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 80.9%; Score 1031; DB 21; Length 259;
 Best Local Similarity 78.7%; Pred. No. 6e-96; 23; Indels 0; Gaps 0;
 Matches 177; Conservative 25; Mismatches 23;

QY 3 GGNOSGHAFFYGGGDSAGMGACGYNLYSGGYGTNTVALSTALFNNGISCGACFEMTC 62
 ||| : |||||
 Db 34 GGNNMAHATFTYGGDSAGTMGACGYNLYSGGYGTNTVALSTALFNNGISCGACFETRC 93
 ||| : |||||
 QY 63 TNDPKMLPCTIRVTATNCPNPFALPNDGGMCPLOHFPMAEPALQIAOYRAGTIVP 122
 || ||||| : |||||
 Db 94 ENDGKWCCLPGSLIVTATNCPNPNMALNNGGCMCPLEHEDLADPVPQRTAOYRAGTIVP 153
 || ||||| : |||||
 QY 123 VSFRRVPCMKKGGVFTTNGHSTFNLVLTITNVGAGDVHSVINGSRTGWSMSNMNGON 182

DB 154 VSYRRVPCRRRGJIEFTTNGSHYFNLYLTITVWGAGDVHSAAIKGSRTVQAMSRMGN 213
QY 183 WSNNTYNGGSLFPVTLTAVLVPNSMOCOTYEGPOF 227
DB 214 WOSNLYNGALSFRTVTSDDRIVSFNAPAGMSYCOTFAGGF 258

RESULT 11
AAG51631
ID AAG51631 standard; Protein: 259 AA.
XX AAG51631;
AC AAG51631;
XX 18-OCT-2000 (first entry)
DT 18-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 65549.
DE Arabidopsis thaliana protein fragment SEQ ID NO: 65549.
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
XX
XX EPI033405-A2.
PN
XX 06-SEP-2000.
PD
XX 25-FEB-2000; 2000EP-0301439.
PF
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 05-MAY-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136382.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139889.
PR 22-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 23-JUN-1999; 99US-0140695.
PR 24-JUN-1999; 99US-0140823.
PR 28-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142134.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146389.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.

```

PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151067.
PR 30-AUG-1999; 99US-0151309.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161933.
PR 29-OCT-1999; 99US-0162142.

```

Query Match

Best Local Similarity 80.9%; Score 1031; DB 21; Length 259;
 Matches 177; Conservative 25; Mismatches 23; Indels 0; Gaps 0;

```

QY 3 GGMOSGATFFYGGGDSGMMGACGCGNIXSGYGTNTVALSTALFNNGLSCGACFEWTC 62
DB 34 GGMNHAATFFYGGGDSGMMGACGCGNLYSGYGTSTAALSTALFNNGLSCGSCFEIRC 93

```

```

QY 63 TNDPKWCLPGTIRVATNCPNPFPALPNDGGMCPLOHPDMAPAFIQTAYRACIVP 122
DB 94 ENDGKWCCLPGSLIVTATNCPNPFPALPNDGGMCPLEHEDLAOPVORTAOTRACIVP 153
QY 123 VSTRRPCKKCKGCVRTTNGHSHYFNLVLTNVGAGDVHVSIGSRGRTGMSRNNCON 182
DB 154 VSTRRPCKKCKGCVRTTNGHSHYFNLVLTNVGAGDVHVSIGSRGRTGMSRNNCON 213
QY 183 WOSNNYLNCGSLFQVLTSDGRTLTAYNLVPSNMQFGQTYEGPOF 227
DB 214 WOSNNYLNCGSLFQVLTSDGRTLTAYNLVPSNMQFGQTYEGPOF 258
RESULT 12
AAG51645
ID AAG51645 standard; Protein: 280 AA.
AC AAG51645;
XX
XX 18-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 65568.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130891.
XX 30-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 04-MAY-1999; 99US-0132407.
XX 05-MAY-1999; 99US-0132484.
XX 06-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 07-MAY-1999; 99US-0132487.
XX 11-MAY-1999; 99US-0132863.
XX 14-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134258.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.

```

PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 19-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145092.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.

PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148654.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161982.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 80.9%; Score 1031; DB 21; Length 280;
Best Local Similarity 78.7%; Pred. No. 6,7e-96;
Matches 177; Conservative 25; Mismatches 23; Indels 0; Gaps 0;

```

OY 3 GGMOSHATFYGGDASGTMGACGYNLVSGYGTNTVALSTALFNNGLSGCAFEPMC 62
DB 55 GGMVNAHATFYGGDASGTMGACGYNLYSGYGTNTALSTALFNNGLSGCAFEPMC 114
OY 63 TNDPKWCLPCTIRVATATNFCPPNFPALPNDGCMCNPPLOHFDMAEPALQIAOYRAGIYP 122
DB 115 ENDGKWCPLGSIYVATATNFCPPNFPALPNDGCMCNPPLOHFDMAEPALQIAOYRAGIYP 174
OY 123 VSFRRVPCMKKGVREFTINGHSYFNVLITNVGAGDVHSVSIKSRFGMOSMRNMGON 182
DB 175 VSTRVPCRRGIRFTINGHSYFNVLITNVGAGDVHSVSIKSRFGMOSMRNMGON 234
OY 183 WQSNNTLVNGGSLSFQVLTSDGRTLTAYNLVPSNMOPQOTYEGPQF 227
DB 235 WQSNNTLVNGGSLSFQVLTSDGRTLTAYNLVPSNMOPQOTYEGPQF 279

RESULT 13
AAR94528
ID AAR94528 standard; Protein: 228 AA.
AC AAR94528;
XX
XX 08-JUL-1996 (first entry)
DE Rice expansin.
KW Expansin; plant cell wall; cellulose; paper recycling; de-linking;
KW polysaccharide; rice.
XX
XX Oryza sativa.
OS
FH Key Location/Qualifiers
FT Misc-difference 211
FT /note= "unidentified amino acid"
PN A09540262-A.
XX
XX 04-APR-1996.
PD
XX 12-MAY-1994; 94AU-0068320.
PF
XX 12-MAY-1995; 95US-0440517.
PR 12-MAY-1993; 93US-0060944.
XX
XX (PENN-) PENN STATE RES FOUND.
PA
XX Cosgrove DJ, McQueen-Mason S;
PI WPI: 1996-201150/21.
XX
XX Expansin proteins which alter the mechanical strength of
XX polysaccharide(s) - useful in paper mfr. and recycling
XX Disclosure; Page 30-31; 60pp; English.
XX
XX Expansins are a novel class of proteins that catalyze the extension
XX of plant cell walls and the weakening of the hydrogen bonds in pure
XX cellulose. 2 Expansins (AAR94528 and AAR94529) have been identified in
XX rice and 3 in Arabidopsis (AAR94530-332). A cDNA clone (AAT13320)
XX coding for cucumber expansin 29 (AAR94527) has been obtd. Expansins
XX can be used e.g. in the mfr. de-linking and recycling of paper, in
XX the textile industry, to aid delignification processes, to alter gel
XX mechanical strength, etc.
XX
XX Sequence 228 AA:
XX
Query Match 80.1%; Score 1021; DB 17; Length 228;
Best Local Similarity 78.8%; Pred. No. 5; 2e-95;
Matches 175; Conservative 23; Mismatches 24; Indels 0; Gaps 0;
OY 3 GGMOSHATFYGGDASGTMGACGYNLVSGYGTNTVALSTALFNNGLSGCAFEPMC 62

```

```

DB 3 GGMVNAHATFYGGDASGTMGACGYNLVSGYGTNTALSTALFNNGLSGCAFEPMC 62
OY 63 TNDPKWCLPCTIRVATATNFCPPNFPALPNDGCMCNPPLOHFDMAEPALQIAOYRAGIYP 122
DB 63 QNDGKWCPLGSIYVATATNFCPPNFPALPNDGCMCNPPLOHFDMAEPALQIAOYRAGIYP 174
OY 123 VSFRRVPCMKKGVREFTINGHSYFNVLITNVGAGDVHSVSIKSRFGMOSMRNMGON 182
DB 123 VSTRVPCRRGIRFTINGHSYFNVLITNVGAGDVHSVSIKSRFGMOSMRNMGON 234
OY 183 WQSNNTLVNGGSLSFQVLTSDGRTLTAYNLVPSNMOPQOTYEGPQF 224
DB 183 WQSNNTLVNGGSLSFQVLTSDGRTLTAYNLVPSNMOPQOTYEGPQF 279

RESULT 14
AAG23852
ID AAG23852 standard; Protein: 251 AA.
AC AAG23852;
XX
XX 17-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 27311.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
OS
PN EP1033405-A2.
XX
XX 06-SEP-2000.
PD
XX 25-FEB-2000; 2000EP-0301439.
PF
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134376.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.

```

PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138544.
PR 14-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139493.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143524.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 21-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.

PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148655.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 20-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156599.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160776.
PR 21-OCT-1999; 99US-0160815.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match	79.9%	Score 1019:	DB 21:	Length 251:
Best Local Similarity	77.7%:	Pred. 10.9e-95:		
Matches 174:	Conservative 24:	Mismatches 25:	Indels 0:	Gaps 0:
QY	4	GMQSHATFYGGDASGTMGACGAGYGNLYSQGTGTNTALSTALEFNNNGISGACFEMTCT	63	
Db	27	GWVNHATFYGGSDSGTMGACGAGYGNLYSQGTGTNTALSTALEFNNNGISGACFEMTCT	63	
QY	64	NDPKACLGCTTATVTATNCNCPNFAFPNDQGGCNPPLOHFDMAEPAFLQIAQYRAGIYVP	123	
Db	87	SDGAMCFLGATITVTATNCNPPNNAAGGACNPPLIHFDSIQVFQRIADYKKGAVVP	146	
QY	124	SFRPAPCKKKGAVRETTINGHSYFNLVLTNNAGCADVSVSIKGSRTQMSKRWGNGNM	183	
Db	147	SYRVPYCMKRCGIRFTTINCHSYFNLVLTNNAGCADVSVAVKGSRTQMSKRWGNGNM	206	
QY	184	QSNNVLLNCGSLSPQVLTSDGRTITAYNLVPSMWQCGQYRECPQF	227	
Db	207	QSNMLNCGALSEFKVTASDGRITVVSNNITAPASMSFGQFTFGRPQF	250	

```

RESULT 15
AAG43343
ID      AAG43343 standard; Protein; 251 AA
XX

```

AC	AA643343;
XX	
DT	18-OCT-2000 (first entry)
XY	

XX Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter termination sequence.

05 Arabidopsis thaliana.
xy

PN Ep1033405-A2

PD 06-SEP-2000.

25-FEB-2000; 2000EP-0301439

PR 25-FEB-1999; 99US-0121825.

PR 09-MAR-1999; 99US-0123548.

PR 25-MAR-1999; 99US-0126264.

PR 01-APR-1999; 99US-0127462.

PR 08-APR-1999; 99US-0128714

PR 19-APR-1999; 99US-0130077

23-APR-1999: 99US-0130510

28-APR-1999: 99US-0137449

30-APR-1999: 0911S-013240Z

05-MAY-1999: 9905-0132484.

06-MAY-1999: 99US-0132486.

99US-0132863.
99US-0134355.
11-MAY-1999;
V/-MAY-1999;
PR

PR 14-MAY-1999; 9905-0134218.

14-MAY-1999; 99US-0134221.
PR 1A-MAY-1999; 99US-0134370

18-MAY-1999; 9905-0134768.
19-MAY-1999; 9905-0134943.
PR

PR 20-MAY-1999; 99US-0135124.

PR	21-MAY-1999;	99SUS-0135535
PR	21-MAY-1999;	99SUS-0135629
PR	25-MAY-1999;	99SUS-0136021
PR	27-MAY-1999;	99SUS-0136392
PR	28-MAY-1999;	99SUS-0136782
PR	01-JUN-1999;	99SUS-0137222
PR	03-JUN-1999;	99SUS-0137528
PR	04-JUN-1999;	99SUS-0137502
PR	07-JUN-1999;	99SUS-0137724
PR	08-JUN-1999;	99SUS-0138094
PR	10-JUN-1999;	99SUS-0138540
PR	10-JUN-1999;	99SUS-0138847
PR	14-JUN-1999;	99SUS-0139119
PR	16-JUN-1999;	99SUS-0139452
PR	17-JUN-1999;	99SUS-0139453
PR	17-JUN-1999;	99SUS-0139492
PR	18-JUN-1999;	99SUS-0139454
PR	18-JUN-1999;	99SUS-0139446
PR	18-JUN-1999;	99SUS-0139455
PR	18-JUN-1999;	99SUS-0139456
PR	18-JUN-1999;	99SUS-0139463
PR	18-JUN-1999;	99SUS-0139457
PR	18-JUN-1999;	99SUS-0139458
PR	18-JUN-1999;	99SUS-0139459
PR	18-JUN-1999;	99SUS-0139460
PR	21-JUN-1999;	99SUS-0139817
PR	22-JUN-1999;	99SUS-0140039
PR	23-JUN-1999;	99SUS-0140353
PR	23-JUN-1999;	99SUS-0140695
PR	24-JUN-1999;	99SUS-0140354
PR	28-JUN-1999;	99SUS-0140823
PR	28-JUN-1999;	99SUS-0140991
PR	30-JUN-1999;	99SUS-0141287
PR	01-JUL-1999;	99SUS-0141842
PR	01-JUL-1999;	99SUS-0142154
PR	02-JUL-1999;	99SUS-0142305
PR	06-JUL-1999;	99SUS-0142399
PR	08-JUL-1999;	99SUS-0142803
PR	12-JUL-1999;	99SUS-0142920
PR	13-JUL-1999;	99SUS-0142977
PR	14-JUL-1999;	99SUS-0143542
PR	15-JUL-1999;	99SUS-0143624
PR	16-JUL-1999;	99SUS-0144005
PR	16-JUL-1999;	99SUS-0144085
PR	19-JUL-1999;	99SUS-0144086
PR	20-JUL-1999;	99SUS-0144332
PR	20-JUL-1999;	99SUS-0144335
PR	20-JUL-1999;	99SUS-0144352
PR	20-JUL-1999;	99SUS-0144632
PR	22-JUL-1999;	99SUS-0144884
PR	21-JUL-1999;	99SUS-0144814
PR	21-JUL-1999;	99SUS-0145086
PR	21-JUL-1999;	99SUS-0145088
PR	22-JUL-1999;	99SUS-0145087
PR	22-JUL-1999;	99SUS-0145085
PR	22-JUL-1999;	99SUS-0145087
PR	22-JUL-1999;	99SUS-0145087
PR	23-JUL-1999;	99SUS-0145115
PR	23-JUL-1999;	99SUS-0145218
PR	23-JUL-1999;	99SUS-0145218
PR	26-JUL-1999;	99SUS-0145224
PR	26-JUL-1999;	99SUS-0145226
PR	27-JUL-1999;	99SUS-0145913
PR	27-JUL-1999;	99SUS-0145918
PR	27-JUL-1999;	99SUS-0145919
PR	28-JUL-1999;	99SUS-0145921
PR	02-AUG-1999;	99SUS-0146586

PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0149368.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 13-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.

PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 79.9%; Score 1019; DB 21; Length 251;
Best Local Similarity 77.7%; Pred. No. 9, 4e-95;
Matches 174; Conservative 24; Mismatches 26; Indels 0; Gaps 0;

QY 4 GMSGHATFYGGDASGTMGACGYGMLYSGQYGTNTVALSTALPFNNGISCGACFEMTCT 63
DB 27 GVNNAHATFYGGSDASGTMGACGYGMLYSGQYGTNTALSTALPFNNGISCGACFEIKQ 86
QY 64 NDPKWCLEPRTIRVATNFCPPNFALPNDGWCNPLQHFDMAPFLQIAYRAGIVPV 123
DB 87 SDGAWCLPGLIIVTATNCPNNALPNNAGWCNPLHFFDLQPVFQRIAYRAGIVPV 146
QY 124 SFRVPCMKKGGVREFTINGHSYFNLVLTIVNGAGDVHSIKGSRGMSMRNNCGNN 183
DB 147 SYRVPCKMRGGIRFTINGHSYFNLVLTIVNGAGDVHSVAVKGSRTRWOOMSRNNCGNN 206
QY 184 OSNNYLNCGLSFOYVTS DGRTLAVNLVPSNMQFGQTEGPOF 227
DB 207 OSNNLNGQALSFKVYASDGRIVSNNTAPASMSFGQTFGRCP 250

Search completed: October 11, 2002, 14:56:45
Job time : 26.1292 secs

RESULT 7
ID MPOL_ORYSA STANDARD: PRT: 263 AA.
AC 040638;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Major pollen allergen Ory s 1 precursor (Ory s 1).
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriharoidae; Oryzaeae; Oryza.
OX NCBI_Taxid=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Another: PubMed-7590339;
RX MEDLINE=96069591; PubMed-7590339;
RA Xu H., Theerakulpisut P., Goultong N., Suphloglu C., Singh M.B.,
RA Bhalla P.L.;
RT "Cloning, expression and immunological characterization of Ory s 1,
RT the major allergen of rice pollen.";
RL Gene 164:255-259(1995).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN MATURE ANTHELS BUT NOT IN
CC VEGETATIVE OR OTHER FLORAL TISSUES.
CC -1- DISEASE: CAUSES GRASS POLLEN ALLERGY.
CC -1- SIMILARITY: BELONGS TO THE LOP P 1 FAMILY OF ALLERGENS.
CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE EG45 DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE CBD DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U31771; AAA86533.1; -
CC HSP: P43214; IMHO.
DR InterPro: IPR000882; Pollen_allergen.
DR Pfam: PF01357; Pollen_allergen.1.
DR PRINTS: PR01225; EXPANSINFAMILY.
DR ProDom: PD002179; Pollen_allergen.1.
DR PROSITE: PS50843; EXPANSIN_CBD.1.
DR PROSITE: PS50842; EXPANSIN_EG45.1.
DR PROSITE: PS50842; EXPANSIN_EG45.1.
KW Allergen; Glycoprotein; Signal; Multigene family.
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 263 MAJOR POLLEN ALLERGEN Ory s 1.
FT DOMAIN 61 164 EXPANSIN-LIKE EG45.
FT DOMAIN 178 259 EXPANSIN-LIKE CBD.
FT CARBOHYD 32 32 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SEQUENCE 263 AA: 28497 MW; B1C5F24EA398DD60 CRC64;
SQ
Query Match 17.4%; Score 216.5; DB 1; Length 263;
Best Local Similarity 27.4%; Pred. No. 2.3e-12;
Matches 55; Conservative 39; Mismatches 82; Indels 25; Gaps 7;

Db 228 PFSVATTEGARRSSAEDAIP 248
RESULT 8
ID GUNS_TIRE STANDARD: PRT: 242 AA.
AC P43317;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Endoglucanase V precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase V)
DE (Cellulase V) (EG V).
GN Egl5.
OS Trichoderma reesei (Hypocrea jecorina).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreales; Hypocreaceae; Hypocrea.
OX NCBI_Taxid=51453;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=QM9414 / RUT C-30;
RX MEDLINE=95075308; PubMed-7984103;
RA Saloheimo A., Henttinen B., Hoffren A.-M., Teeman O., Penttilae M.;
RA "A novel, small endoglucanase gene, egl5, from Trichoderma reesei
RT isolated by expression in yeast.";
RT Mol. Microbiol. 13:219-228(1994).
CC -1- CAPTIVIT ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
CC linkages in cellulose.
CC -1- SIMILARITY: BELONGS TO CELLULOSE FAMILY K (FAMILY 45 OF GLYCOSYL
CC HYDROLASES).
CC -1- SIMILARITY: CONTAINS 1 FUNGAL-TYPE CELLULOSE-BINDING DOMAIN (CBD).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: Z3381; GAA83846.1; -
CC HSP: P00725; 2CBH.
DR InterPro: IPR000254; CBD_fungal.
DR InterPro: IPR000334; CBD_hydro_45.
DR Pfam: PF00734; CBD_1.1.
DR Pfam: PF02015; Glyco_hydro_45.1.
DR ProDom: PD001821; CBD_fungal.1.
DR SMART: SM00236; fcbd.1.
DR PROSITE: PS00562; CBD_FUNGAL.1.
DR PROSITE: PS50842; EXPANSIN_EG45.1.
DR PROSITE: PS01140; GLYCOSYL_HYDROL_F45.1.
DR PROSITE: PS01140; GLYCOSYL_HYDROL_F45.1.
KW Cellulose degradation; Hydrolase; Glycosidase; Signal.
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 242 ENDOGLUCANASE V.
FT ACT SITE 27 27 NUCLEOPHILE (BY SIMILARITY).
FT ACT SITE 134 134 PROTON DONOR (BY SIMILARITY).
FT DOMAIN 18 182 CATALYTIC.
FT DOMAIN 183 205 PRO/SER-RICH (LINKER).
FT DOMAIN 206 242 CELLULOSE-BINDING (BY SIMILARITY).
FT CARBOHYD 182 182 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 213 230 BY SIMILARITY.
FT DISULFID 224 240 BY SIMILARITY.
FT SEQUENCE 242 AA: 24411 MW; CC033F51326C71D CRC64;
SQ
Query Match 8.0%; Score 99; DB 1; Length 242;
Best Local Similarity 30.4%; Pred. No. 0.067;
Matches 38; Conservative 11; Mismatches 28; Indels 48; Gaps 9;

RESULT 5
 MP21_CYNDA
 ID MP21_CYNDA STANDARD: PRT: 246 AA.
 AC 004701;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Major pollen allergen Cyn d 1.
 DE CYNDA.
 OS Cynodon dactylon (Bermuda grass).
 OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta;
 OC Spermatophyta: Magnoliophyta: Liliopsida: Poales; Poaceae; PACC clade;
 OC Chloridoideae; Cynodonteae; Cynodon.
 RX NCBI_TaxID=28909;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Pollen;
 RX MEDLINE=96347957; PubMed=8757211;
 RA Smith P.M., Suphloglu C., Griffith I.J., Thierault K., Knox R.B.,
 RA Singh M.B.;
 RT Cloning and expression in yeast *Pichia pastoris* of a biologically
 RT active form of Cyn d 1, the major allergen of Bermuda grass pollen."
 RL J. Allergy Clin. Immunol. 98:331-343(1996).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- DISEASE: CAUSES GRASS POLLEN ALLERGY.
 CC -1- SIMILARITY: BELONGS TO THE LOI P I FAMILY OF ALLERGENS.
 CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE EG45 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE CBD DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: S83343; AAB50734.2; -
 DR HSSP: P43214; IMHO.
 DR InterPro: IPR000882; Pollen_allergen.
 DR Pfam: PF01357; Pollen_allergen.1.
 DR PRINTS: PRO1225; EXPANSINFAMILY.
 DR ProDom: PD002179; Pollen_allergen.1.
 DR PROSITE: PS50843; EXPANSIN_CBD; 1.
 DR PROSITE: PS50842; EXPANSIN_EG45; 1.
 DR Allergen: 39 145 EXPANSIN-LIKE EG45.
 FT DOMAIN 159 240 EXPANSIN-LIKE CBD.
 FT CARBOHYD 9 9 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 246 AA; 26888 MW; 43D8442DBA58322 CRC64;
 Query Match 20.0%; Score 249.5; DB 1; Length 246;
 Best Local Similarity 29.2%; Pred. No. 2.4e-15;
 Matches 70; Conservative 41; Mismatches 90; Indels 39; Gaps 9;
 QY 5 WERGHAFYVG---GADASGTMGAGCYGNLHSGYGLQTAALSTALFNSGKCGACPEL 60
 DB 18 WLEARAFYGSNPRGA-APDDHGACGYKDYDPRPDGMTACGNEIFPDGIGCRACEI 76
 QY 61 TCEDDEPWCIPGISTIVRYNLANFALANDNGMGNPPLKHFDLAEPFLQIAO-----Y 113
 DB 77 KCK-EPVCGEGEPLVITDKNE-----HIAHYFDLSGKAFGMAKKGQEDKIR 126
 QY 114 RAGIVPAFRVRVCEKGGGIRFTIN---GNPFDLVLITNWGAGADIRAVLSKSKTDQ 169
 DB 127 KAGELTQIFRRVCKYPSGKTKTFHEKGSNDHYALLLVKVAAGCNIVAVDKPRDSE 186
 QY 170 WQSHRNMGOMQSNRY--LRGSLSFQVYTDSPGRTVVSVDVPHDM-----QPG 217
 DB 187 FIRMKSMGAIHRIIDPKPLKG-PFSIRLTSGGALHVGDDVIFANWKPDIYVYSKIQFG 245

RESULT 6
 MP21_MAIZE
 ID MP21_MAIZE STANDARD: PRT: 191 AA.
 AC 007154;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Pollen allergen Zea m 1 (Zea m 1).
 DE Zea mays (Maize).
 OS Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta;
 OC Spermatophyta: Magnoliophyta: Liliopsida: Poales; Poaceae; PACC clade;
 OC Panicoideae; Andropogoneae; Zea.
 RX NCBI_TaxID=4577;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pollen;
 RX MEDLINE=94010312; PubMed=8406014;
 RA Broadwater A.H., Rubinstein A.L., Chay C.H., Klapper D.G.,
 RA Redinger P.A.;
 RT "Zea m1, the maize homolog of the allergen-encoding loi p1 gene of
 RT rye grass."
 RL Gene 131:227-230(1993).
 CC -1- TISSUE SPECIFICITY: POLLEN TISSUE.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSION LOW BEFORE AND HIGH AFTER
 CC POLLEN MITOSIS.
 CC -1- DISEASE: CAUSES MAIZE POLLEN ALLERGY.
 CC -1- SIMILARITY: BELONGS TO THE LOI P I FAMILY OF ALLERGENS.
 CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE EG45 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE CBD DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: L14271; AAA3496.1; -
 DR PIR: JC1524; JC1524.
 DR HSSP: P43214; IMHO.
 DR MaizEdB: 65840; -
 DR InterPro: IPR000882; Pollen_allergen.
 DR Pfam: PF01357; Pollen_allergen.1.
 DR PRINTS: PRO1225; EXPANSINFAMILY.
 DR ProDom: PD002179; Pollen_allergen.1.
 DR PROSITE: PS50843; EXPANSIN_CBD; 1.
 DR PROSITE: PS50842; EXPANSIN_EG45; 1.
 DR Allergen: Multigene family.
 FT DOMAIN 1 91 EXPANSIN-LIKE EG45.
 FT DOMAIN 105 186 EXPANSIN-LIKE CBD.
 SQ SEQUENCE 191 AA; 21362 MW; 6E2A9DF921C45C63 CRC64;
 Query Match 17.7%; Score 220; DB 1; Length 191;
 Best Local Similarity 28.3%; Pred. No. 7.9e-13;
 Matches 54; Conservative 39; Mismatches 70; Indels 28; Gaps 7;
 QY 40 TAASTALFNSGKCGACPELTCEDDEPWCIPGISTIVRYNLANFALANDNGMGNPPLK- 98
 DB 2 TACGNVPIFDGKGCSCYEVRCERKEPE-CSGNPVIVFTDMNY-----EPLAP 49
 QY 99 -HFDLAEPFLQIAO-----YRAGIVPAFRVRVCEKGGGIRFTIN---GNPFDLV 146
 DB 50 YHPLSGKAFGSLAKGGLNDKLHRCGIMDEFFRRVCKYPAQCKIVFHEKGCNPNYAV 109
 QY 147 LITNWGAGDIRAVLSKSKTDQMSNRNMGOMQSNRY--YLRGSLSFQVYTDSPGRTV 204
 DB 110 LVEFVADDDGIVLMEIQDKLSAEWKPMKLSGAINRMOTAKLKG-PFSIRLTSGSKV 168
 QY 205 VSYDVVPHDMQ 215
 DB 169 IAKDIIPANMR 179

```

CC EMBL: S80654; AAB35984.1; -.
DR HSSP: P43214; IMHO.
DR InterPro: IPR000882; Pollen_allergen.
DR Pfam: PF01357; Pollen_allergen.1.
DR PRINTS: PR01225; EXPANSINFAMILY.
DR ProDom: PD002179; Pollen_allergen.1.
DR PROSITE: PS50843; EXPANSIN_CBD; 1.
DR PROSITE: PS50842; EXPANSIN_EG45; 1.
KW Allergen; Glycoprotein; Signal.
FT SIGNAL 1 29
FT CHAIN 1 29
FT DOMAIN 67 173 EXPANSIN-LIKE EG45.
FT DOMAIN 187 268 EXPANSIN-LIKE CBD.
FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 37 37 P -> G (IN REF. 2).
SQ SEQUENCE 269 AA; 29011 MW; 058834458ACE877F CRC64;

Query Match 20.6%; Score 257; DB 1; Length 269;
Best Local Similarity 30.7%; Pred. No. 5.6e-16;
Matches 70; Conservative 35; Mismatches 93; Indels 30; Gaps 9;

OY 5 WERGHATFTYGADASGTM--GGACGYNLHSGYGLQTAALSTALFNSQKCGACFELTC 62
DB 47 WLDKASTWYKPTGAGPKDNGAGCYKVDKAPFNGMTGCGNTPFKRGGSGSEFLKC 106
OY 63 EDDPENCIPGSIIVRYNLNLFALANDNGCMNPPLK--HFDLAEPALQIAO-----Y 113
DB 107 -SKPESGSGEATVH-----ITDDN-----EPIIAPYHFDLSGHAFGSMARKGEENVR 154
OY 114 RAGIVPAFRFRPCEKGGGIRFTIN--GNPFDLVLTNNGAGDIRAVLSKSKTDQ 169
DB 155 GAGLELEQFRVRYKVKYDPTKPFVHEKSGSNPVYLAIVKYVDGDDVYAVDIREKGD 214
OY 170 WQSMRNWQGNQMSNT--YLKGSLSFOVTDSDGRVVSVDVYPHDMQ 215
DB 215 WIELKESGAIWIRIDTPDKLTG--PFTVRYTTEGGTKAEFDVIREGK 261

RESULT 4
MPLI_LOLPR STANDARD; PRT; 263 AA.
AC P14946; P1964;
ID MPLI_LOLPR
DT 01-APR-1990 (Rel. 14, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Pollen allergen Lol p 1 precursor (Lol p 1) (Allergen R7).
OS Lolium perenne (Perennial ryegrass).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
OC Poaceae; Lolium.
OX NCBI_TaxID=4522;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90375479; PubMed=1697854;
RA Perez M., Ishioka G.Y., Walker L.E., Chesnut R.W.;
RT "CDNA cloning and immunological characterization of the rye grass
RT allergen Lol p 1.";
RL J. Biol. Chem. 265:16210-16215(1990).
RN [2]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 24-48.
RX TISSUE=Pollen;
RA MEDLINE=91160716; PubMed=2001733;
RA Griffith I.J., Smith P.M., Pollock J., Theerakulpisut P.,
RA Avjoglou A., Davies S., Hough T., Singh M.B., Simpson R.J., Ward L.D.,
RA Knox R.B.;
RT "Cloning and sequencing of Lol p1, the major allergenic protein of
RT rye-grass pollen.";
RL FEBS Lett. 279:210-215(1991).
RN [3]
RP SEQUENCE OF 24-53.
RX TISSUE=Pollen;
RN MEDLINE=86242068; PubMed=3718469;

```

```

RA Cottam G.P., Moran D.M., Standring R.;
RT "Physicochemical and immunochemical characterization of allergenic
RT proteins from rye-grass (Lolium perenne) pollen prepared by a rapid
RT and efficient purification method.";
RL Biochem. J. 234:305-310(1986).
RN [4]
RP SEQUENCE OF 236-263.
RX MEDLINE=89364850; PubMed=2475768;
RA Esch R.E., Klapper D.G.;
RT "Isolation and characterization of a major cross-reactive grass group
RT allergenic determinant.";
RL Mol. Immunol. 26:557-561(1989).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DISEASE: CAUSES GRASS POLLEN ALLERGY.
CC -1- SIMILARITY: BELONGS TO THE LOL P 1 FAMILY OF ALLERGENS.
CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE EG45 DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE CBD DOMAIN.
CC -----
CC "THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M57474; AAA63279.1; -.
DR EMBL: M57476; AAA63278.1; -.
DR PIR: A23341; A23341.
DR PIR: B37881; B37881.
DR PIR: S13614; S13614.
DR HSSP: P43214; IMHO.
DR InterPro: IPR000882; Pollen_allergen.
DR Pfam: PF01357; Pollen_allergen.1.
DR PRINTS: PR01225; EXPANSINFAMILY.
DR ProDom: PD002179; Pollen_allergen.1.
DR PROSITE: PS50843; EXPANSIN_CBD; 1.
DR PROSITE: PS50842; EXPANSIN_EG45; 1.
KW Allergen; Glycoprotein; Signal; Multigene family.
FT SIGNAL 1 23
FT CHAIN 24 263
FT DOMAIN 61 167 EXPANSIN-LIKE EG45.
FT DOMAIN 181 262 EXPANSIN-LIKE CBD.
FT CARBOHYD 32 32 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 68 68 N -> D.
FT VARIANT 177 177 A -> G.
FT VARIANT 210 210 I -> T.
FT VARIANT 246 246 F -> V.
FT VARIANT 28 28 P -> C (IN REF. 3).
FT CONFLICT 31 31 P -> C (IN REF. 3).
FT CONFLICT 48 48 W -> MW (IN REF. 3).
SQ SEQUENCE 263 AA; 28438 MW; 7675896F279C88C9 CRC64;

Query Match 20.4%; Score 254; DB 1; Length 263;
Best Local Similarity 29.8%; Pred. No. 1e-15;
Matches 68; Conservative 37; Mismatches 93; Indels 30; Gaps 9;

OY 5 WERGHATFTYGADASGTM--GGACGYNLHSGYGLQTAALSTALFNSQKCGACFELTC 62
DB 41 WLDKASTWYKPTGAGPKDNGAGCYKVDKAPFNGMTGCGNTPFKRGGSGSEFLKC 100
OY 63 EDDPENCIPGSIIVRYNLNLFALANDNGCMNPPLK--HFDLAEPALQIAO-----Y 113
DB 101 -TPKESGSGEATVH-----ITDDN-----EPIIAPYHFDLSGHAFGSMARKGEENVR 148
OY 114 RAGIVPAFRFRPCEKGGGIRFTIN--GNPFDLVLTNNGAGDIRAVLSKSKTDQ 169
DB 149 GAGLELEQFRVRYKVKYDPTKPFVHEKSGSNPVYLAIVKYVDGDDVYAVDIREKGD 208
OY 170 WQSMRNWQGNQMSNT--YLKGSLSFOVTDSDGRVVSVDVYPHDMQ 215
DB 209 WIELKESGAIWIRIDTPDKLTG--PFTVRYTTEGGTKAEFDVIREGK 255

```

```

OY 5 WERGHATFYGADASGTM--GCACGYGNLHSGYGLDTALSTALFNSGKGCACFELTC 62
DB 41 WLDKSTWYKPGKAGKNDGACGKDYDKPPSGMTGCGNPIFRSGGSCSEFELIC 100
OY 63 EDDPENCIPGSIIVRYNLANFALANDNGWCNPLK--HFDLAEPFLQIAQ-----Y 113
DB 101 -TRFACSGEPYVH-----ITDDN-----EPIAPYHFDLSGHAFGAMAKKGEQKLR 148
OY 114 RAGIVPAFRVRPCEKGGIRFTIN---GNPYFDLVLTNNGAGDIRAVSLKSKRTDQ 169
DB 149 SAGLELQFRVRCCKYKPEFHEKGSNPYLLALVYDGDVAVADIKKEDK 208
OY 170 WQSMRNMGMQMSNT--YLRGSLSFQYTDSDGRTVSVYDVVPHDMPGQFTE 221
DB 209 WIELKESWGMVWVDPDKLTG--PFTVRYTTGCTGAEADVIEGKADTSTE 261

RESULT 2
MPAL_HOLLA STANDARD: PRT; 265 AA.
AC P43216: Q39975;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Major pollen allergen Hol 1 precursor (Hol 1 I) (Hol 1 I.0101 and
DE 1.0102).
OS Holcus lanatus (Velvet grass).
OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
OC Spermatophyta: Magnoliophyta: Liliopsida: Poales: Poaceae: Poideae:
OC Poaceae: Holcus.
OX NCBI_TaxID=29679;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. AVENEAE; TISSUE=Pollen;
RA Schramm G.D., Bufe A., Becker W.M., Schlaak M.;
RT Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 18-265 FROM N.A.
RC STRAIN=CV. AVENEAE; TISSUE=Pollen;
RA Schramm G.D., Bufe A., Petersen A., Haas H., Schlaak M., Becker W.M.;
RT Mapping of IgE-binding epitopes on the recombinant major group I
RT allergen of velvet grass pollen, rHol 1 I.*;
RL J. Allergy Clin. Immunol. 99:781-787(1997).
RN [3]
RP CHARACTERIZATION.
RC STRAIN=CV. AVENEAE; TISSUE=Pollen;
RA MEDLINE=96319506; PubMed=8768803;
RA Schramm G.D., Petersen A., Bufe A., Schlaak M., Becker W.M.;
RT Identification and characterization of the major allergens of velvet
RT grass (Holcus lanatus), Hol 1 I and Hol 1.5.*;
RL Int. Arch. Allergy Immunol. 110:354-363(1996).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE LOL P I FAMILY OF ALLERGENS.
CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE EG45 DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE CBD DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z27084; CAAB1610.1; -
DR EMBL: Z68893; CAAG3121.1; -
DR HSSP: P43214; 1WHO.
DR InterPro: IPR000882; Pollen.allergen.
DR Pfam: PF01357; Pollen.allergen.1.
DR PRINTS: PR01225; EXPANSINFAMILY.
DR ProDom: PD002179; Pollen.allergen.1.

```

```

DR PROSITE: PS50843; EXPANSIN_CBD; 1.
DR PROSITE: PS50842; EXPANSIN_EG45; 1.
KW Allergen; Signal.
FT SIGNAL 1 25
FT CHAIN 26 265
FT DOMAIN 63 169 MAJOR POLLEN ALLERGEN HOL 1.1.
FT DOMAIN 183 264 EXPANSIN-LIKE EG45.
FT VARIANT 103 103 EXPANSIN-LIKE CBD.
FT VARIANT 103 103 T -> S (IN HOL 1.1.0102).
SQ SEQUENCE 265 AA; 28590 MW; 1FE23B3FE198AD6D CRC64;

Query Match
Best Local Similarity 29.1%; Pred. No. 5,5e-16;
Matches 68; Conservative 40; Mismatches 96; Indels 30; Gaps 9;

OY 5 WERGHATFYGADASGTM--GCACGYGNLHSGYGLDTALSTALFNSGKGCACFELTC 62
DB 43 WLDKSTWYKPGKAGKNDGACGKDYDKPPSGMTGCGNPIFRSGGSCSEFELIC 102
OY 63 EDDPENCIPGSIIVRYNLANFALANDNGWCNPLK--HFDLAEPFLQIAQ-----Y 113
DB 103 -TRFACSGEPYVH-----ITDDN-----EPIAPYHFDLSGHAFGAMAKKGEQKLR 150
OY 114 RAGIVPAFRVRPCEKGGIRFTIN---GNPYFDLVLTNNGAGDIRAVSLKSKRTDQ 169
DB 151 SAGLELQFRVRCCKYKPEFHEKGSNPYLLALVYDGDVAVADIKKEDK 210
OY 170 WQSMRNMGMQMSNT--YLRGSLSFQYTDSDGRTVSVYDVVPHDMPGQFTE 221
DB 211 WIELKESWGMVWVDPDKLTG--PFTVRYTTGCTGAEADVIEGKADTSTE 263

RESULT 3
MPAL_PHAQO STANDARD: PRT; 269 AA.
AC Q41260;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Major pollen allergen Pha a 1 precursor (Pha a I).
OS Phalaris aquatica (Canary grass).
OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
OC Spermatophyta: Magnoliophyta: Liliopsida: Poales: Poaceae: Poideae:
OC Poaceae: Phalaris.
OX NCBI_TaxID=28479;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pollen;
RA MEDLINE=96105569; PubMed=8564724;
RA Supphlogl C., Singh M.B.;
RT Cloning, sequencing and expression in Escherichia coli of pha a 1
RT and four isoforms of pha a 5, the major allergens of canary grass
RT pollen.*;
RL Clin. Exp. Allergy 25:853-865(1995).
RN [2]
RP SEQUENCE OF 30-49.
RC MEDLINE=93319091; PubMed=7687099;
RA Supphlogl C., Singh M.B., Simpson R.J., Ward L., Knox R.B.;
RT Identification of canary grass (Phalaris aquatica) pollen allergens
RT by immunoblotting: IgE and IgG antibody-binding studies.*;
RL Allergy 48:273-281(1993).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DISEASE: CAUSES GRASS POLLEN ALLERGY.
CC -1- SIMILARITY: BELONGS TO THE LOL P I FAMILY OF ALLERGENS.
CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE EG45 DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE CBD DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 11, 2002, 14:50:30 : Search time 5.8181 Seconds
(without alignments)
1499.000 Million cell updates/sec

Title: US-09-896-301-5

Sequence: 1 DNGGWERGHAFYGCADASG.....SYDVVPHDWQFGQTFEGGQF 225

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	272	21.8	263	1	MPPL_PHLPR
2	257	20.6	265	1	MPPL_HOLLA
3	257	20.6	265	1	MPPL_PHAQ
4	254	20.4	263	1	MPPL_LOLPR
5	249.5	20.0	246	1	MPCL_CYNDA
6	220	17.7	191	1	MPPL_MAIZE
7	216.5	17.4	263	1	MPOL_ORVSA
8	99	8.0	242	1	GUN5_TRIRE
9	97.5	7.8	333	1	CATC_MOUSE
10	89	7.1	462	1	GUN_MOUSE
11	87.5	7.0	181	1	GUN_MYTED
12	85.5	6.9	97	1	MP2_LOLPR
13	85.5	6.9	1250	1	YFAL_ECOLI
14	83	6.7	463	1	CATC_HUMAN
15	81	6.5	418	1	GUN2_TRIRE
16	81	6.5	878	1	FTMD_ECOLI
17	80.5	6.5	1744	1	RPAL_TRYBB
18	79.5	6.4	997	1	ATST_HUMAN
19	79	6.3	564	1	ESTJ_HELVY
20	79	6.3	870	1	FTMD_HELVY
21	79	6.3	1192	1	IMG2_MOUSE
22	78.5	6.3	373	1	ADHX_RABIT
23	78	6.3	1545	1	IGA3_HAELN
24	77.5	6.2	316	1	Y007_CHLTR
25	77.5	6.2	326	1	VS09_ROTRO
26	77.5	6.2	1592	1	SORL_CHICK
27	77	6.2	386	1	RB87_DROME
28	77	6.2	462	1	CATC_RAT
29	77	6.2	2334	1	WAPA_BACSU
30	76.5	6.1	771	1	SKNI_YEAST
31	76	6.1	466	1	HTOA_HAELN
32	76	6.1	1172	1	LMB3_HUMAN
33	75.5	6.1	378	1	ELYA_BACSP

34	75	6.0	254	1	YVNC_BACSU	006977 bacillus su
35	75	6.0	486	1	GICA_THEMA	09wz27 thermotoga
36	74.5	6.0	173	1	FLAV_HAELN	P44562 haemophilus
37	74.5	6.0	570	1	STIM_DROME	P83094 drosophila
38	74.5	6.0	755	1	TRZM_AGR4	P04029 agrobacteri
39	74.5	6.0	798	1	HMBP_DROME	P31264 drosophila
40	74.5	6.0	1541	1	ICAI_HAELN	P42782 haemophilus
41	74	5.9	482	1	SYP_METTH	026708 methanobact
42	74	5.9	1222	1	PMP5P_HUMAN	Q15155 homo sapien
43	74	5.9	2327	1	CCAB_MOUSE	055017 mus musculu
44	74	5.9	2336	1	CCAB_RAT	Q02294 rattus norv
45	74	5.9	2339	1	CCAB_HUMAN	Q00975 homo sapien

ALIGNMENTS

RESULT 1
ID MPPL_PHLPR STANDARD: PRT: 263 AA.
AC P43213:
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Pollen allergen Phl p 1 precursor (Phl p 1).
GN PHLP.
OS Phleum pratense (Common timothy).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Poaceae; Phleum.
OX NCBI_TaxID=15957;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pollen;
RX MEDLINE=95015525; PubMed=7930302;
RA Laffer S., Valente R., Vrtala S., Susani M., van Ree R., Kraft D.,
RA Scheiner O., Duchene M.;
RT "Complementary DNA cloning of the major allergen Phl p 1 from timothy
grass (Phleum pratense); recombinant Phl p 1 inhibits IgE binding to
group I allergens from eight different grass species.";
RT J. Allergy Clin. Immunol. 94:689-698(1994).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE LOI p 1 FAMILY OF ALLERGENS.
CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE EG45 DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE CBD DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X78813; CA55390.1; -.
CC HSSP: P43214; IMHO.
DR InterPro: IPR000882; Pollen.allergen.
DR Pfam: PF01357; Pollen.allergen; 1.
DR PRINTS: PR01225; EXPANSINFAM1.
DR PROSITE: PD002179; Pollen.allergen; 1.
DR PROSITE: PS50843; EXPANSIN_CBD; 1.
DR PROSITE: PS50842; EXPANSIN_EG45; 1.
KW Allergen; Glycoprotein; Signal; Multigene family.
FT SIGNAL 1 23
FT CHAIN 1 263
FT DOMAIN 61 167
FT DOMAIN 181 262
FT CARBOHYD 32 32
SQ SEQUENCE 263 AA; 28457 MW; 046BA249C17BC048 CRC64;
Query Match 21.88; Score 272; DB 1; Length 263;
Best Local Similarity 29.9%; Pred. No. 2.5e-17;
Matches 70; Conservative 41; Mismatches 93; Indels 30; Gaps 9;

DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
 DE EXPANSIN.
 GN EXP2.
 OS Prunus armeniaca (Apricot).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.
 OX NCBI_TaxID=36596;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERGERON; TISSUE=MESOCARP, EXOCARP.
 RA Mdeguite-A-Mdeguite D., Fils-Lycaon B.;
 RT "Molecular cloning and nucleotide sequence of expansin 2 (Pa-Exp2)
 RT from apricot fruit."
 RL Submitted (DEC-1997) to the EMBL/genbank/DBJ databases.
 DR EMBL: AF038815; AAC33530.1;
 DR InterPro: IPR000882; Pollen_allergen.
 DR Pfam: PF01357; Pollen_allergen; 1.
 DR PRINTS: PR01225; EXPANSNFAMILY.
 DR Prodom: PD002179; Pollen_allergen; 1.
 SO SEQUENCE 252 AA; 26758 MW; 5B81AE5C8E959791 CRC64;

Query Match 76.7%; Score 954.5; DB 10; Length 252;
 Best Local Similarity 76.5%; Pred. No. 3.4e-81;
 Matches 173; Conservative 21; Mismatches 27; Indels 5; Gaps 3;

OY 3 GGEHGATFTYGGADASGTMGACGYGNLSOGYGLQTAALSTALFNSGKCGACFELTC 62
 DB 29 GGEHGATFTYGGADASGTMGACGYGNLSOGYGLQTAALSTALFNSGKCGACFELTC 88
 OY 63 EDDPEWICPGSIIV--RYNLANFALANDNGWCNPPLEKHFDLAPFLQIAQYRAGIIV 119
 DB 89 DSDPKWCLPGSIIVATNCPNLAQSNNDNGWCNPPLEKHFDLAPFLQIAQYRAGIIV 148
 OY 120 VAFRRVPCCKGGIRFTTNGNPFYDLVLTNNVGAGDITRAVSLKSKTDQMSRMNGQ 179
 DB 149 ISRRVSCVKKGGIRFTTNGNPFYDLVLTNNVGAGDVHSVSIKSKTG-WQMSRMNGQ 207
 OY 180 NMOSNTYLRGOSLSPQVTDSDGRVTSYDVYPHDMQFQTFEGGQF 225
 DB 208 NMOSNTYLRGOSLSPQVTDSDGRVTS-NAVYADMQFQTFEGGQF 252

RESULT 11
 OYFS30 PRELIMINARY; PRT; 252 AA.
 AC OYFS30:
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
 DE EXPANSIN.
 GN PPEXP1.
 OS Prunus persica (Peach).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.
 OX NCBI_TaxID=3760;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Hayama H.;
 RA "Homolog to expansin in peach fruit."
 RT Submitted (JUN-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AB029083; BAB19676.1;
 DR InterPro: IPR000882; Pollen_allergen.
 DR PRINTS: PR01225; EXPANSNFAMILY.
 DR Prodom: PD002179; Pollen_allergen; 1.
 SO SEQUENCE 252 AA; 26743 MW; BF0B86FE0C060482 CRC64;

Query Match 76.3%; Score 950.5; DB 10; Length 252;
 Best Local Similarity 76.5%; Pred. No. 8.1e-81;
 Matches 173; Conservative 21; Mismatches 27; Indels 5; Gaps 3;

OY 3 GGEHGATFTYGGADASGTMGACGYGNLSOGYGLQTAALSTALFNSGKCGACFELTC 62

DB 29 GGEHGATFTYGGADASGTMGACGYGNLSOGYGLQTAALSTALFNSGKCGACFELTC 88
 OY 63 EDDPEWICPGSIIV--RYNLANFALANDNGWCNPPLEKHFDLAPFLQIAQYRAGIIV 119
 DB 89 DSDPKWCLPGSIIVATNCPNLAQSNNDNGWCNPPLEKHFDLAPFLQIAQYRAGIIV 148
 OY 120 VAFRRVPCCKGGIRFTTNGNPFYDLVLTNNVGAGDITRAVSLKSKTDQMSRMNGQ 179
 DB 149 ISRRVSCVKKGGIRFTTNGNPFYDLVLTNNVGAGDVHSVSIKSKTG-WQMSRMNGQ 207
 OY 180 NMOSNTYLRGOSLSPQVTDSDGRVTSYDVYPHDMQFQTFEGGQF 225
 DB 208 NMOSNTYLRGOSLSPQVTDSDGRVTS-NAVYADMQFQTFEGGQF 252

RESULT 12
 OYFS11 PRELIMINARY; PRT; 253 AA.
 AC OYFS11:
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
 DE EXPANSIN.
 GN EXP2.
 OS Fragaria ananassa (Strawberry).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids I; Rosales; Rosaceae; Rosoideae; Fragaria.
 OX NCBI_TaxID=3747;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. CHANDLER.
 RA Civello P.M., Sabehat A., Powell A.L.T., Bennett A.B.;
 RT "An expansin gene expressed in ripening strawberry fruit is auxin-
 RT independent."
 RL Plant Physiol. 12:1273-1279(1999).
 DR EMBL: AF159563; AAF21101.1;
 DR InterPro: IPR000882; Pollen_allergen.
 DR Pfam: PF01357; Pollen_allergen; 1.
 DR PRINTS: PR01225; EXPANSNFAMILY.
 DR Prodom: PD002179; Pollen_allergen; 1.
 DR PROSITE: PS00626; RC01.2; UNKNOWN.1.
 SO SEQUENCE 253 AA; 26887 MW; CE39CF00ADEA1CEF CRC64;

Query Match 75.8%; Score 944; DB 10; Length 253;
 Best Local Similarity 76.4%; Pred. No. 3.3e-80;
 Matches 172; Conservative 18; Mismatches 31; Indels 4; Gaps 2;

OY 4 GGEHGATFTYGGADASGTMGACGYGNLSOGYGLQTAALSTALFNSGKCGACFELTC 63
 DB 30 GGEHGATFTYGGADASGTMGACGYGNLSOGYGLQTAALSTALFNSGKCGACFELTC 89
 OY 64 DDEPWCIPGSIIV--RYNLANFALANDNGWCNPPLEKHFDLAPFLQIAQYRAGIIV 120
 DB 90 NDPRWCLPGSIIVATNCPNLAQSNNDNGWCNPPLEKHFDLAPFLQIAQYRAGIIV 149
 OY 121 AFRVPCCKGGIRFTTNGNPFYDLVLTNNVGAGDITRAVSLKSKTDQMSRMNGQ 180
 DB 150 SFRVPCCKGGIRFTTNGNPFYDLVLTNNVGAGDVHSVSIKSKTG-WQMSRMNGQ 208
 OY 181 WOSTYLRGOSLSPQVTDSDGRVTSYDVYPHDMQFQTFEGGQF 225
 DB 209 WOSTYLRGOSLSPQVTDSDGRVTS-NAVYADMQFQTFEGGQF 253

RESULT 13
 OYFS493 PRELIMINARY; PRT; 232 AA.
 AC OYFS493:
 DT 01-MAY-1997 (TREMBLrel. 03, Created)
 DT 01-MAY-1997 (TREMBLrel. 03, last sequence update)

RT extension in plants.
 RL Proc. Natl. Acad. Sci. U.S.A. 92:9245-9249(1995).
 DR EMBL: U30382; AAB37746.1;
 DR InterPro: IPR000882; Pollen_allergen.
 DR Pfam: PF01357; Pollen_allergen.1.
 DR PRINTS: PR01225; EXPANSNFAMLY.
 DR Prodom: PD002179; Pollen_allergen.1.
 RT CHAIN 24 250 EXPANSIN SL.
 SQ SEQUENCE 250 AA; 27215 MW; 60651BC47EA186DA CRC64;

Query Match
 Best Local Similarity 77.3%; Score 962; DB 10; Length 250;
 Matches 171; Conservative 26; Mismatches 27; Indels 4; Gaps 2;

QY 1 DNGGMRGHATFYGGADASGTMGACGCGNLSGQYGLQTAALSTALFNSGCGACFEL 60
 DB 24 DYGMQSHATFYGGADASGTMGACGCGNLSGQYGLQTAALSTALFNSGCGACFEL 83
 QY 61 TCEDDPEMCIPGSIIV---RYNLANFALANDGMCNPLKHFDAEPALQIAQYRA 117
 DB 84 TCTNDPRKMCIPGSIIVRYNLATNFCPPNFPALPNNNGMCNPLQHFDAEPALQIAQYRA 143
 QY 118 VPAFAFRVPCCKGIRFTINGNPFYDLVLTNNGACDRAVSLKSKTDQMSMRW 177
 DB 144 VPAFAFRVPCCKGIRFTINGNPFYDLVLTNNGACDRAVSLKSKTDQMSMRW 202
 QY 178 GGNMOSNTYLRGQSLSFQVTDGRTVSYDVPHDMQFGTFEGGQF 225
 DB 203 GGNMOSNTYLRGQSLSFQVTDGRTVSYDVPHDMQFGTFEGGQF 250

RESULT 8
 ID 022874 PRELIMINARY; PRT: 253 AA.

AC 022874
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE PUTATIVE EXPANSIN.
 GN AT2G0610.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_Taxid=3702;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RX MEDLINE=20083487; PubMed=10617197;
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
 RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
 RA Cronin L.A., Shen M., Vanaken S.E., Umayam L., Tallon L.J., Gill J.E.,
 RA Adams M.D., Gaitera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
 RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
 RA Salzberg S.L., Fraser C.M., Venter J.C.;
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
 thaliana."
 RL Nature 402:761-768(1999).
 RN 12
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RX LIn X.;
 RL EMBL: AC002336; AAB87577.1;
 DR InterPro: IPR000882; Pollen_allergen.
 DR Pfam: PF01357; Pollen_allergen.1.
 DR PRINTS: PR01225; EXPANSNFAMLY.
 DR Prodom: PD002179; Pollen_allergen.1.
 SQ SEQUENCE 253 AA; 27260 MW; 08C42860D523DD8A CRC64;

Query Match
 Best Local Similarity 77.0%; Score 958.5; DB 10; Length 253;
 Matches 171; Conservative 26; Mismatches 27; Indels 4; Gaps 2;

Matches 171; Conservative 23; Mismatches 29; Indels 7; Gaps 2;

QY 1 DNGGMRGHATFYGGADASGTMGACGCGNLSGQYGLQTAALSTALFNSGCGACFEL 60
 DB 26 DYGMQSHATFYGGADASGTMGACGCGNLSGQYGLQTAALSTALFNSGCGACFEL 85
 QY 61 TCEDDPEMCIPGSIIVRYNLANF-----ALANDGMCNPLKHFDAEPALQIAQYRA 115
 DB 86 KCNDPRKMCIPGSIIVRYNLATNFCPPNFPALPNNNGMCNPLQHFDAEPALQIAQYRA 143
 QY 116 GIVPAFAFRVPCCKGIRFTINGNPFYDLVLTNNGACDRAVSLKSKTDQMSMR 175
 DB 144 GIVPAFAFRVPCCKGIRFTINGNPFYDLVLTNNGACDRAVSLKSKTDQMSMR 203
 QY 176 GGNMOSNTYLRGQSLSFQVTDGRTVSYDVPHDMQFGTFEGGQF 225
 DB 204 GGNMOSNTYLRGQSLSFQVTDGRTVSYDVPHDMQFGTFEGGQF 253

RESULT 9
 ID 09FNT0 PRELIMINARY; PRT: 260 AA.

AC 09FNT0
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE EXPANSIN.
 OS Cicer arietinum (chickpea) (Garbanzo).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Ciceraceae; Cicer.
 OX NCBI_Taxid=3827;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. CASTELLANA; TISSUE=ETIOLATED EPICOTYLS;
 RX Dopico B., Sanchez M.A., Labrador E.;
 RT "An second expansin is expressed in chickpea epicotyls."
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ291817; CAC19184.1;
 DR InterPro: IPR000882; Pollen_allergen.
 DR Prodom: PD002179; Pollen_allergen.1.
 DR Prosite: PS00030; RRM_RNP-1; UNKNOWN.1.
 SQ SEQUENCE 260 AA; 28266 MW; 32A70368ED2883E9 CRC64;

Query Match
 Best Local Similarity 76.9%; Score 957; DB 10; Length 260;
 Matches 172; Conservative 23; Mismatches 27; Indels 8; Gaps 3;

QY 1 DNGGMRGHATFYGGADASGTMGACGCGNLSGQYGLQTAALSTALFNSGCGACFEL 60
 DB 34 DYGMQSHATFYGGADASGTMGACGCGNLSGQYGLQTAALSTALFNSGCGACFEL 93
 QY 61 TCEDDPEMCIPGSIIVRYNLANF-----ALANDGMCNPLKHFDAEPALQIAQYRA 115
 DB 94 KCNDPRKMCIPGSIIVRYNLATNFCPPNFPALPNNNGMCNPLQHFDAEPALQIAQYRA 151
 QY 116 GIVPAFAFRVPCCKGIRFTINGNPFYDLVLTNNGACDRAVSLKSKTDQMSMR 175
 DB 152 GIVPAFAFRVPCCKGIRFTINGNPFYDLVLTNNGACDRAVSLKSKTDQMSMR 210
 QY 176 GGNMOSNTYLRGQSLSFQVTDGRTVSYDVPHDMQFGTFEGGQF 225
 DB 211 GGNMOSNTYLRGQSLSFQVTDGRTVSYDVPHDMQFGTFEGGQF 260

RESULT 10
 ID 081133 PRELIMINARY; PRT: 252 AA.
 AC 081133
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)

OY 1 DNGMGRHATFEYGGADASCTMGACGYNLHSGYGLQTAALSTALFNSSGCKGACFEL 60
 DB 28 DNGMGRHATFEYGGADASCTMGACGYNLHSGYGLQTAALSTALFNSSGCKGACFEL 87
 OY 61 TCEDDPEMCIPGSIIV---RYNLANFALANDGMCNPPKHPDLAEPALQIAOYRAGI 117
 DB 88 TCEDDPEMCIPGSIIVSATNFCPPNFALANDGMCNPPKHPDLAEPALQIAOYRAGI 147
 OY 118 VPAVAFRRVPCCKGGIRFTINGNPFYDLVLTINVGAGADIRAVSLKSGKTDQMSMRNW 177
 DB 148 VPAVAFRRVPCCKGGIRFTINGNPFYDLVLTINVGAGADIRAVSLKSGKTDQMSMRNW 207
 OY 178 GQNMOSNTYLKRGOSLSFOVTDSDGRTVVSVDVPHDMQFGQTFEGGQF 225
 DB 208 GQNMOSNTYLKRGOSLSFOVTDSDGRTVVSVDVPHDMQFGQTFEGGQF 255

RESULT 2

O9FLC5 PRELIMINARY: PRT: 255 AA.
 AC O9FLC5: 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE EXPANSIN AF-EXP2
 OS Arabidopsis thaliana (mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eustosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=COLUMBIA;
 RX MEDLINE=96344145; PubMed=9679202;
 RA Kaneko T., Kotani H., Nakamura Y., Sato S., Asamizu E., Miyajima N.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. V. Sequence
 RT features of the regions of 1,381,565 bp covered by twenty one
 RT physically assigned P1 and TMC clones."
 RL DNA Res. 5:131-143(1996).
 DR EMBL: AB010692; BAR09972.1; -;
 DR InterPro: IPR000882; Pollen_allergen.
 DR Pfam: PF01357; Pollen_allergen; 1.
 DR PRINTS: PR01225; EXPANSINFAMILY.
 DR Prodom: PD002179; Pollen_allergen; 1.
 SO SEQUENCE 255 AA; 27749 MW; 53E678633B2C4152 CRC64;

Query Match 96.2%; Score 1197.5; DB 10; Length 255;
 Best Local Similarity 96.1%; Pred. No. 7.4e-104;
 Matches 219; Conservative 1; Mismatches 5; Indels 3; Gaps 1;

OY 1 DNGMGRHATFEYGGADASCTMGACGYNLHSGYGLQTAALSTALFNSSGCKGACFEL 60
 DB 28 DNGMGRHATFEYGGADASCTMGACGYNLHSGYGLQTAALSTALFNSSGCKGACFEL 87
 OY 61 TCEDDPEMCIPGSIIV---RYNLANFALANDGMCNPPKHPDLAEPALQIAOYRAGI 117
 DB 88 TCEDDPEMCIPGSIIVSATNFCPPNFALANDGMCNPPKHPDLAEPALQIAOYRAGI 147
 OY 118 VPAVAFRRVPCCKGGIRFTINGNPFYDLVLTINVGAGADIRAVSLKSGKTDQMSMRNW 177
 DB 148 VPAVAFRRVPCCKGGIRFTINGNPFYDLVLTINVGAGADIRAVSLKSGKTDQMSMRNW 207
 OY 178 GQNMOSNTYLKRGOSLSFOVTDSDGRTVVSVDVPHDMQFGQTFEGGQF 225
 DB 208 GQNMOSNTYLKRGOSLSFOVTDSDGRTVVSVDVPHDMQFGQTFEGGQF 255

RESULT 3

O9LNB2 PRELIMINARY: PRT: 245 AA.
 AC O9LNB2: 01-OCT-2000 (Tremblrel. 15, Created)

DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE EXPANSIN 2.
 OS Zinnia elegans.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids II; Asterales; Asteraceae; Asteroideae;
 OC Helianthaceae; Zinnia.
 OX NCBI_TaxID=34245;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=20317189; PubMed=10859177;
 RX Im K.H., Cosgrove D.J., Jones A.M.;
 RT "Subcellular localization of expansin mRNA in xylem cells."
 RL Plant Physiol. 123:463-470(2000).
 DR EMBL: AF230332; AAF35901.1; -;
 DR InterPro: IPR000882; Pollen_allergen.
 DR Pfam: PF01357; Pollen_allergen; 1.
 DR PRINTS: PR01225; EXPANSINFAMILY.
 DR Prodom: PD002179; Pollen_allergen; 1.
 SO SEQUENCE 245 AA; 26201 MW; F66E26114C27E360 CRC64;

Query Match 78.9%; Score 982; DB 10; Length 245;
 Best Local Similarity 75.9%; Pred. No. 9e-84;
 Matches 173; Conservative 25; Mismatches 22; Indels 8; Gaps 3;

OY 3 GQMGHATFEYGGADASCTMGACGYNLHSGYGLQTAALSTALFNSSGCKGACFEL 62
 DB 21 GQMGHATFEYGGADASCTMGACGYNLHSGYGLQTAALSTALFNSSGCKGACFEL 80
 OY 63 EDDPEMCIPGSIIVRYNLANFALANDGMCNPPKHPDLAEPALQIAOYRAGI 117
 DB 81 NDDPKMCLPGSIIV---TATNFCPPNFGSLNDGMCNPPKHPDLAEPALQIAOYRAGI 138
 OY 118 VPAVAFRRVPCCKGGIRFTINGNPFYDLVLTINVGAGADIRAVSLKSGKTDQMSMRNW 177
 DB 139 VPISTFQVPCVKKGVYFTINGSHYENLITINVGAGADIRAVSLKSGKTDQMSMRNW 197
 OY 178 GQNMOSNTYLKRGOSLSFOVTDSDGRTVVSVDVPHDMQFGQTFEGGQF 225
 DB 198 GQNMOSNTYLKRGOSLSFOVTDSDGRTVVSVDVPHDMQFGQTFEGGQF 245

RESULT 4

O9FUM3 PRELIMINARY: PRT: 254 AA.
 AC O9FUM3: 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE EXPANSIN 1.
 OS Prunus avium (Cherry), and
 OS Prunus cerasus.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC euroids I; Rosales; Rosaceae; Amygdaloideae; Prunus.
 OX NCBI_TaxID=42229, 140311;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=P. avium;
 RA Wu Z., Wiersma P.A.;
 RT "Differential Expression of Expansin Genes Isolated from Sweet Cherry
 RT (Prunus avium L.) during Fruit Ripening."
 RL Submitted (Aug-2000) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=P. cerasus; TISSUE=RIPE FRUIT;
 RA Yoo S.-D., Gao Z., Cantini C., Loescher W., van Nocker S.;
 RT "Coordinated expression of genes encoding expansins and other cell
 RT wall-modifying enzymes is associated with pectin-related changes in
 RT the cell wall during ripening of cherry (P. cerasus) fruit."
 RL Submitted (Feb-2001) to the EMBL/Genbank/DBJ databases.

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 11, 2002, 14:51:00 ; Search time 19.428 seconds
(without alignments)
2003.488 Million cell updates/sec

Title: US-09-896-301-5

Perfect score: 1245

Sequence: 1 DNGGMRGHATFGGADASG.....SYDVPHDMQFGQTFFEGGQF 225

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

SPREMBL_19:*

- 1: sp.archaea:*
- 2: sp.bacteria:*
- 3: sp.fungi:*
- 4: sp.human:*
- 5: sp.invertebrate:*
- 6: sp.mammal:*
- 7: sp.mhc:*
- 8: sp.organelle:*
- 9: sp.phage:*
- 10: sp.plant:*
- 11: sp.rodent:*
- 12: sp.virus:*
- 13: sp.vertebrate:*
- 14: sp.unclassified:*
- 15: sp.rvivirus:*
- 16: sp.bacteriaph:*
- 17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1203.5	96.7	255	10	Q38866
2	1197.5	96.2	255	10	Q9FLC5
3	982	78.9	245	10	Q9LUB2
4	972	78.1	254	10	Q9FUM3
5	970	77.9	254	10	Q82093
6	963	77.3	253	10	Q93KP2
7	962	77.3	250	10	Q93625
8	958.5	77.0	253	10	Q22874
9	957	76.9	260	10	Q9FMT0
10	954.5	76.7	252	10	Q81133
11	950.5	76.3	252	10	Q9FES30
12	944	75.8	253	10	Q9SRT1
13	927	74.5	232	10	P93493
14	927	74.5	232	10	P93495
15	927	74.5	253	10	Q9SWY1
16	923	74.1	232	10	P93492

17	919.5	73.9	253	10	Q9SMD4
18	912	73.3	232	10	P93494
19	912	73.3	249	10	Q9LDR9
20	912	73.3	258	10	Q49194
21	905	72.7	247	10	Q82625
22	901	72.4	249	10	Q92P35
23	898	72.1	251	10	Q946J1
24	886	71.2	247	10	Q9M517
25	884	71.0	246	10	P93442
26	883.5	71.0	242	10	Q9LUB1
27	883	70.9	246	10	Q946J0
28	883	70.9	251	10	Q40636
29	876	70.4	249	10	Q93KP1
30	875	70.3	252	10	Q9FTM2
31	871	70.0	253	10	Q94KT7
32	869	69.8	237	10	Q38863
33	869	69.8	250	10	Q9CS54
34	868	69.7	248	10	Q80622
35	866.5	69.6	249	10	Q9M515
36	858	68.9	255	10	Q9FMA0
37	851	68.4	249	10	Q9ZP36
38	851	68.4	257	10	Q48818
39	850	68.3	262	10	Q80932
40	849	68.2	239	10	Q9ZP31
41	846	68.0	250	10	Q9FVG9
42	846	68.0	260	10	Q9M2S9
43	844	67.8	258	10	Q41043
44	840	67.5	260	10	Q93XP0
45	834	67.0	257	10	Q9XG16

ALIGNMENTS

RESULT 1

ID Q38866 PRELIMINARY: PRT: 255 AA.

AC Q38866;

DT 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE EXPANSIN AT-EXP2.

GN AT-EXP2.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Kossidae;

OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.

OX NCBI_TaxID=3702;

ON [1]

RN SEQUENCE FROM N.A.

RP MEDLINE-96016146; PubMed-7568110;

RX Shecherban T.Y., Shi J., Durachko D.M., Gulltman M.J.,

RA McQueen-Mason S.J., Shieh M., Cosgrove D.J.;

RT "Molecular cloning and sequence analysis of expansins--a highly

RT conserved, multigene family of proteins that mediate cell wall

RT extension in plants.";

RT Proc. Natl. Acad. Sci. U.S.A. 92:9245-9249(1995).

RL [2]

RN SEQUENCE FROM N.A.

RP Shecherban T.Y., Shi J., Durachko D.M., Gulltman M.J.,

RA McQueen-Mason S.J., Shieh M., Cosgrove D.J.;

RL Submitted (MAR-2001) to the EMBL/GenBank/DBD databases.

DR EMBL, U30481; AAB38073.1; -

DR InterPro: IPR000882; Pollen.allergen.

DR Pfam: PF01357; Pollen.allergen; 1.

DR PRINTS: PR01225; EXPANSINFAMILY.

DR PRODOM: PD002179; Pollen.allergen; 1.

SQ SEQUENCE 255 AA; 2785F0827A285500 CRC64;

Query Match 96.7%; Score 1203.5; DB 10; Length 255;

Best Local Similarity 96.5%; Pred. No. 2e-104;

Matches 220; Conservative 1; Mismatches 4; Indels 3; Gaps 1;

PR 14-MAY-1999: 99US-0134370.
PR 18-MAY-1999: 99US-0134476.
PR 19-MAY-1999: 99US-0134941.
PR 20-MAY-1999: 99US-0135124.
PR 21-MAY-1999: 99US-0135353.
PR 24-MAY-1999: 99US-0135629.
PR 25-MAY-1999: 99US-0136021.
PR 27-MAY-1999: 99US-0136392.
PR 28-MAY-1999: 99US-0136782.
PR 01-JUN-1999: 99US-0137222.
PR 03-JUN-1999: 99US-0137528.
PR 04-JUN-1999: 99US-0137502.
PR 07-JUN-1999: 99US-0137724.
PR 08-JUN-1999: 99US-0138094.
PR 10-JUN-1999: 99US-0138540.
PR 10-JUN-1999: 99US-0138847.
PR 14-JUN-1999: 99US-0139119.
PR 16-JUN-1999: 99US-0139452.
PR 17-JUN-1999: 99US-0139492.
PR 18-JUN-1999: 99US-0139454.
PR 18-JUN-1999: 99US-0139455.
PR 18-JUN-1999: 99US-0139456.
PR 18-JUN-1999: 99US-0139457.
PR 18-JUN-1999: 99US-0139458.
PR 18-JUN-1999: 99US-0139459.
PR 18-JUN-1999: 99US-0139460.
PR 18-JUN-1999: 99US-0139461.
PR 18-JUN-1999: 99US-0139462.
PR 18-JUN-1999: 99US-0139463.
PR 18-JUN-1999: 99US-0139750.
PR 18-JUN-1999: 99US-0139763.
PR 21-JUN-1999: 99US-0139817.
PR 22-JUN-1999: 99US-0139899.
PR 23-JUN-1999: 99US-0140353.
PR 23-JUN-1999: 99US-0140354.
PR 24-JUN-1999: 99US-0140695.
PR 28-JUN-1999: 99US-0140823.
PR 29-JUN-1999: 99US-0140891.
PR 30-JUN-1999: 99US-0141287.
PR 01-JUL-1999: 99US-0141842.
PR 01-JUL-1999: 99US-0142154.
PR 02-JUL-1999: 99US-0142055.
PR 06-JUL-1999: 99US-0142390.
PR 08-JUL-1999: 99US-0142803.
PR 09-JUL-1999: 99US-0142820.
PR 12-JUL-1999: 99US-0142977.
PR 13-JUL-1999: 99US-0143452.
PR 14-JUL-1999: 99US-0143624.
PR 15-JUL-1999: 99US-0144005.
PR 16-JUL-1999: 99US-0144085.
PR 16-JUL-1999: 99US-0144086.
PR 19-JUL-1999: 99US-0144325.
PR 19-JUL-1999: 99US-0144331.
PR 19-JUL-1999: 99US-0144332.
PR 19-JUL-1999: 99US-0144333.
PR 19-JUL-1999: 99US-0144334.
PR 19-JUL-1999: 99US-0144335.
PR 20-JUL-1999: 99US-0144352.
PR 20-JUL-1999: 99US-0144632.
PR 20-JUL-1999: 99US-0144884.
PR 21-JUL-1999: 99US-0144814.
PR 21-JUL-1999: 99US-0145086.
PR 21-JUL-1999: 99US-0145088.
PR 22-JUL-1999: 99US-0145087.
PR 22-JUL-1999: 99US-0145089.
PR 22-JUL-1999: 99US-0145192.
PR 23-JUL-1999: 99US-0145145.
PR 23-JUL-1999: 99US-0145218.
PR 23-JUL-1999: 99US-0145224.
PR 26-JUL-1999: 99US-0145276.
PR 27-JUL-1999: 99US-0145913.

PR 27-JUL-1999: 99US-0145918.
PR 27-JUL-1999: 99US-0145919.
PR 28-JUL-1999: 99US-0145951.
PR 28-AUG-1999: 99US-0146386.
PR 02-AUG-1999: 99US-0146388.
PR 02-AUG-1999: 99US-0146389.
PR 03-AUG-1999: 99US-0147038.
PR 04-AUG-1999: 99US-0147204.
PR 04-AUG-1999: 99US-0147302.
PR 05-AUG-1999: 99US-0147192.
PR 05-AUG-1999: 99US-0147260.
PR 06-AUG-1999: 99US-0147303.
PR 06-AUG-1999: 99US-0147416.
PR 09-AUG-1999: 99US-0147935.
PR 09-AUG-1999: 99US-0147935.
PR 10-AUG-1999: 99US-0148171.
PR 11-AUG-1999: 99US-0148319.
PR 12-AUG-1999: 99US-0148341.
PR 13-AUG-1999: 99US-0148565.
PR 13-AUG-1999: 99US-0148684.
PR 16-AUG-1999: 99US-0149368.
PR 17-AUG-1999: 99US-0149175.
PR 18-AUG-1999: 99US-0149426.
PR 20-AUG-1999: 99US-0149722.
PR 20-AUG-1999: 99US-0149723.
PR 20-AUG-1999: 99US-0149925.
PR 23-AUG-1999: 99US-0149902.
PR 23-AUG-1999: 99US-0149930.
PR 25-AUG-1999: 99US-0150566.
PR 26-AUG-1999: 99US-0150884.
PR 27-AUG-1999: 99US-0151065.
PR 27-AUG-1999: 99US-0151066.
PR 27-AUG-1999: 99US-0151080.
PR 30-AUG-1999: 99US-0151303.
PR 31-AUG-1999: 99US-0151438.
PR 01-SEP-1999: 99US-0151930.
PR 07-SEP-1999: 99US-0152363.
PR 10-SEP-1999: 99US-0153070.
PR 13-SEP-1999: 99US-0153759.
PR 15-SEP-1999: 99US-0154018.
PR 16-SEP-1999: 99US-0154039.
PR 20-SEP-1999: 99US-0154779.
PR 22-SEP-1999: 99US-0155139.
PR 23-SEP-1999: 99US-0155486.
PR 24-SEP-1999: 99US-0155659.
PR 28-SEP-1999: 99US-0156458.
PR 29-SEP-1999: 99US-0156596.
PR 05-OCT-1999: 99US-0157117.
PR 05-OCT-1999: 99US-0157753.
PR 06-OCT-1999: 99US-0157865.
PR 07-OCT-1999: 99US-0158029.
PR 08-OCT-1999: 99US-0158232.
PR 12-OCT-1999: 99US-0158369.
PR 13-OCT-1999: 99US-0159293.
PR 13-OCT-1999: 99US-0159294.
PR 13-OCT-1999: 99US-0159295.
PR 14-OCT-1999: 99US-0159329.
PR 14-OCT-1999: 99US-0159330.
PR 14-OCT-1999: 99US-0159331.
PR 14-OCT-1999: 99US-0159637.
PR 14-OCT-1999: 99US-0159638.
PR 18-OCT-1999: 99US-0159584.
PR 21-OCT-1999: 99US-0160741.
PR 21-OCT-1999: 99US-0160767.
PR 21-OCT-1999: 99US-0160768.
PR 21-OCT-1999: 99US-0160770.
PR 21-OCT-1999: 99US-0160814.
PR 21-OCT-1999: 99US-0160815.
PR 22-OCT-1999: 99US-0160980.
PR 22-OCT-1999: 99US-0160981.
PR 22-OCT-1999: 99US-0160989.
PR 25-OCT-1999: 99US-0161404.
PR 25-OCT-1999: 99US-0161405.

PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140655.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141287.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142350.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145193.
PR 23-JUL-1999; 99US-0145195.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.

PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match

62.2%; Score 744.5; DB 21; Length 259;


```
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151086.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151920.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156596.
PR 29-SEP-1999; 99US-0157117.
PR 04-OCT-1999; 99US-0157173.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159284.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161997.
PR 29-OCT-1999; 99US-0162142.

Query Match 62.2%; Score 744.5; DB 21; Length 249;
Best Local Similarity 63.4%; Pred. No. 3e-68;
Matches 135; Conservative 27; Mismatches 48; Indels 3; Gaps 2;
```

```
QY 7 SAFAFFYGGKDDSGCTWGCAGCGNLYNMGYGLYNALSSALFNDGACGACTTTCDSO 66
DB 28 NAHAFFYGGGDSAGTWCAGCGNLYSGYGTSTAALSTALEFNNGLSCGSCFEIRCE-ND 86
QY 67 TTKWCKRGSGSITTTATNLCXPNMALPSNCGMCPPLKHFDPKSOPIAMENIAYVQAGIVP 126
DB 87 GKMKCLG--SIVVTATNFCPPNNALANNNGMCNPLEFEDLAQGVFQIAQYRAGIVP 144
QY 127 NYKRVFXQSGGIRFAISCHDYFELVTVTNVGGSGVVAQMSIKSGNTGMMASRMNGANW 186
DB 145 SYRRAVCRRRGGRIRFTINHSHYFNLVLTITVCGAGDVHSAALKSGRTVQASRMNGQNW 204
```

```
QY 187 QSNAYLAGOSLSFTVOLDGKRYTANNAKPPXNM 219
DB 205 QSNSYINGQALSFKRYTSDGRFVSFNAPAGW 237

RESULT 10
ID AAGS1632
ID AAGS1632 standard; Protein; 249 AA.
AC AAGS1632;
XX
XX
XX 18-OCT-2000 (first entry)
XX
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 65550.
XX
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX
XX
XX 06-SEP-2000.
XX
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130891.
XX 30-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 06-MAY-1999; 99US-0132487.
XX 07-MAY-1999; 99US-0132863.
XX 11-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137528.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 16-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139492.
```


PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159337.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 62.2%; Score 744.5; DB 21; Length 241;
Best Local Similarity 63.4%; Pred. No. 2.9e-68;
Matches 135; Conservative 27; Mismatches 48; Indels 3; Gaps 2;

OY 7 SAFATFYGGKDGSCCTMGAGCGYNYNAGCYGLYNALSSALFENDGACGACCTTTCDSQ 66
DB 20 NHAHATFYGGGDSGTMGAGCGYNYNAGCYGLYNALSSALFENDGACGACCTTTCDSQ 78
OY 67 TKWCKPGNSITITATNLCXPMALPNSGWCNPLXHFDMSPAWENIVYQAGIYV 126
DB 79 GKWCPLG--SIVVTATNCPNNALANNNGWCNPLXHFDMSPAWENIVYQAGIYV 136
OY 127 NKRRVXPXRSRGIRFALISGHDFELVTVTVNAGSVVAQMSIKGSNTGMMASRMNGANW 186
DB 137 SYRRVPCRRRGIRFALISGHDFELVTVTVNAGSVVAQMSIKGSNTGMMASRMNGANW 196
OY 187 OSNAVLAGOSLSFTYQDDGGRKVTAMNAPXNW 219
DB 197 OSNSTLNGQALSFKVTYTDGRTVVSFNAAPAGW 229

RESULT 8
ID AAG51647
XX AAG51647 standard; Protein: 241 AA.
AC AAG51647;

XX 18-OCT-2000 (first entry)
DT Arabidopsis thaliana protein fragment SEQ ID NO: 65570.
DE Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
OS EP1033405-A2.
XX 06-SEP-2000.
PD 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136382.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.

KW Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EPI033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130891.
XX 30-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 04-MAY-1999; 99US-0132407.
XX 05-MAY-1999; 99US-0132484.
XX 06-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 06-MAY-1999; 99US-0132487.
XX 07-MAY-1999; 99US-0132863.
XX 11-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137232.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137582.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 16-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139459.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
XX 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.
XX 18-JUN-1999; 99US-0139460.
XX 18-JUN-1999; 99US-0139461.
XX 18-JUN-1999; 99US-0139462.
XX 18-JUN-1999; 99US-0139463.
XX 18-JUN-1999; 99US-0139750.
XX 18-JUN-1999; 99US-0139763.
XX 21-JUN-1999; 99US-0139817.
XX 22-JUN-1999; 99US-0139899.
XX 23-JUN-1999; 99US-0140353.

PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148365.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.

PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142290.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0143472.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145813.
PR 27-JUL-1999; 99US-0145819.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.

PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154479.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match
Best Local Similarity 62.2%; Score 744.5; DB 21; Length 241;
Matches 135; Conservative 27; Mismatches 48; Indels 3; Gaps 2;

QY 7 SAFATFYGGKDCCTMGACGYGNLYMAGYGLYMAALSATFNDGAMGACYTITCDTSQ 66
Db 20 NHAFTYGGGDASGTMGACGYGNLYSQYGTSTALSTALEFNNGLSCSCFEIRCE-ND 78
QY 67 TWCKRBGNSITITATNLCPMMLPNSGCGNCPPLXHFDMSQPAMENIAVYQGIYPV 126
Db 79 GWCCLPG--SIVTATNFPPNNALANNCGNCPPLHFHDLAOFVFORIAQYRAGIYPV 136
QY 127 NKRVXORSRGIRFRAISGHDFELVTYTNVGSQVVAOMSIKGSNTGMMASRWGANW 186
Db 137 STRVPCRRRGIRFTINSHSYFNILVLTITNAGADYSAATKGSRTYVQAMSRWGMW 196
QY 187 QSNAYLAGOSLSFTVQLDDGRKVTAMNXPXNM 219
Db 197 QSNYLYNGALSFTKVTTSQGRTVSEFNAPACM 229

RESULT 7
AAG51633
ID AAG51633 standard; Protein; 241 AA.
XX AAG51633;
AC AAG51633;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 65551.
XX

XX W0200123530-A1.
XX
XX 05-APR-2001.
PD
XX 29-SEP-2000; 2000WO-US26884.
XX
XX 30-SEP-1999; 99US-0410191.
PR
XX (RECC) UNIV CALIFORNIA.
PA
XX Bradford KJ, Chen F, Dahal P, Downie B, Nonogaki H;
PI
XX MPI: 2001-266144/27.
DR
XX N-PSDB: AAD03712.
XX
XX Novel nucleic acid sequences isolated from germinating seeds encoding
PT polypeptides that are useful to control seed germination in plants -
PS
XX Claim 3; Page 49; 63pp; English.
XX
XX The present invention relates to enzymes such as seed-specific
CC endo-beta-mannanase (e.g., Lycopersicon esculentum (Lc) MAN2),
CC polygalacturonase (PG) (e.g., LcPG1), cellulases such as Cel5 and
CC Cel68, arabinosidase (e.g., LcAra), xyloglucan endotransglycosylases
CC (XET) (e.g., LcXET4) and expansins such as LcExp4, LcExp8 and LcExp10
CC isolated from germinating seeds. All these enzymes are expressed
CC initially in the endosperm caps and are associated with cell wall
CC hydrolysis. These enzymes are associated with weakening of tissues
CC surrounding the embryo and/or initiating radicle growth. The control of
CC expression of these endogenous genes is therefore a convenient means for
CC controlling seed germination. The present sequence is tomato seed
CC expansin, LcExp8. Expansins are extracellular proteins that
XX facilitate cell wall extension.
SQ
XX Sequence 257 AA:
Query Match 63.9%; Score 765; DB 22; Length 257;
Best Local Similarity 65.0%; Pred. No. 2,4e-70;
Matches 139; Conservative 21; Mismatches 52; Indels 2; Gaps 2;
OY 7 SAFAFYCGKSSCTMGACGCGYNYMGGLYNAALSSAIFNCGACACTTTCDS- 65
DB 33 SAHAFTYCGADASTGTMGACGCGYNYMGGLYNAALSSAIFNCGACACTTTCDS- 92
OY 66 QTKMCKPGCNSITITATLCLXPNMALPSNSGCMPLXHPDMSQPAWENIAVYQAGIIP 125
DB 93 DPQMKCK-CVSTYITITSTNFCPPNYLPSNNGCMCPRRPHFDMQPAWEKGIKGIIP 151
OY 126 VMYKRVKXQSGGIRFAISGHDYFELVTVTVNGSGVYAQMSIKGNTGMAMSRNMGAN 185
DB 152 VLYKRVKCKKHGCVFTINGRDYFELVTVSNVGGAGSYESVQIKGNTNMWLTMSRMGAS 211
OY 186 WOSNAYLGGOSLFTVOLDGKRVAMXAPXNM 219
DB 212 WOSNAYLGGOSLFTVOLDGKRVAMXAPXNM 245
RESULT 6
AAG06546
ID AAG06546 standard; Protein: 241 AA.
XX
XX AAG06546;
AC
XX
XX 17-OCT-2000 (first entry)
DT
XX
XX Arabidopsis thaliana protein fragment SFO ID NO: 3357.
DE
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
OS

XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
PD
XX
XX 25-FEB-2000; 2000EP-0301439.
PE
XX
XX 25-FEB-1999; 99US-0121825.
PR
XX 05-MAR-1999; 99US-0123180.
PR
XX 09-MAR-1999; 99US-0123548.
PR
XX 23-MAR-1999; 99US-0125788.
PR
XX 25-MAR-1999; 99US-0126264.
PR
XX 29-MAR-1999; 99US-0126785.
PR
XX 01-APR-1999; 99US-0127462.
PR
XX 06-APR-1999; 99US-0128234.
PR
XX 08-APR-1999; 99US-0128714.
PR
XX 16-APR-1999; 99US-0129845.
PR
XX 19-APR-1999; 99US-0130077.
PR
XX 21-APR-1999; 99US-0130449.
PR
XX 23-APR-1999; 99US-0130510.
PR
XX 23-APR-1999; 99US-0130891.
PR
XX 28-APR-1999; 99US-0131449.
PR
XX 30-APR-1999; 99US-0132048.
PR
XX 30-APR-1999; 99US-0132407.
PR
XX 04-MAY-1999; 99US-0132484.
PR
XX 05-MAY-1999; 99US-0132485.
PR
XX 06-MAY-1999; 99US-0132486.
PR
XX 06-MAY-1999; 99US-0132487.
PR
XX 07-MAY-1999; 99US-0132863.
PR
XX 11-MAY-1999; 99US-0134256.
PR
XX 14-MAY-1999; 99US-0134218.
PR
XX 14-MAY-1999; 99US-0134219.
PR
XX 14-MAY-1999; 99US-0134221.
PR
XX 14-MAY-1999; 99US-0134370.
PR
XX 18-MAY-1999; 99US-0134768.
PR
XX 19-MAY-1999; 99US-0134941.
PR
XX 20-MAY-1999; 99US-0135124.
PR
XX 21-MAY-1999; 99US-0135353.
PR
XX 24-MAY-1999; 99US-0135629.
PR
XX 25-MAY-1999; 99US-0136021.
PR
XX 27-MAY-1999; 99US-0136392.
PR
XX 28-MAY-1999; 99US-0136782.
PR
XX 03-JUN-1999; 99US-0137222.
PR
XX 04-JUN-1999; 99US-0137528.
PR
XX 07-JUN-1999; 99US-0137502.
PR
XX 08-JUN-1999; 99US-0137724.
PR
XX 10-JUN-1999; 99US-0138094.
PR
XX 10-JUN-1999; 99US-0138540.
PR
XX 14-JUN-1999; 99US-0138847.
PR
XX 16-JUN-1999; 99US-0139119.
PR
XX 16-JUN-1999; 99US-0139452.
PR
XX 16-JUN-1999; 99US-0139453.
PR
XX 17-JUN-1999; 99US-0139492.
PR
XX 18-JUN-1999; 99US-0139454.
PR
XX 18-JUN-1999; 99US-0139455.
PR
XX 18-JUN-1999; 99US-0139456.
PR
XX 18-JUN-1999; 99US-0139457.
PR
XX 18-JUN-1999; 99US-0139458.
PR
XX 18-JUN-1999; 99US-0139459.
PR
XX 18-JUN-1999; 99US-0139460.
PR
XX 18-JUN-1999; 99US-0139461.
PR
XX 18-JUN-1999; 99US-0139462.
PR
XX 18-JUN-1999; 99US-0139463.
PR
XX 18-JUN-1999; 99US-0139750.
PR
XX 18-JUN-1999; 99US-0139753.
PR
XX 21-JUN-1999; 99US-0139817.
PR
XX 22-JUN-1999; 99US-0139819.
PR
XX 23-JUN-1999; 99US-0140353.
PR
XX 23-JUN-1999; 99US-0140354.
PR
XX 24-JUN-1999; 99US-0140695.
PR
XX 28-JUN-1999; 99US-0140823.
PR
XX 29-JUN-1999; 99US-0140991.
PR
XX 30-JUN-1999; 99US-0141287.

PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140033.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.

PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147302.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152365.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.

XX 12-MAY-1994; 94AU-0068320.
 PF 12-MAY-1995; 95US-0440517.
 PR 12-MAY-1993; 93US-0060944.
 XX
 XX (PENN-) PENN STATE RES FOUND.
 PI Cosgrove DJ, McQueen-Mason S;
 XX WPI: 1996-201150/21.
 DR

XX Expansin proteins which alter the mechanical strength of
 PT poly:saccharide(s) - useful in paper mfr. and recycling
 XX
 XX Disclosure: Page 31-32; 60pp; English.
 PS

CC Expansins are a novel class of proteins that catalyse the extension
 CC of plant cell walls and the weakening of the hydrogen bonds in pure
 CC cellulose. 2 Expansins (AAR94528 and AAR94529) have been identified in
 CC rice and 3 in Arabidopsis (AAR94530-32). A cDNA clone (AAT13320)
 CC coding for cucumber expansin 29 (AAR94527) has been obtd. Expansins
 CC can be used e.g. in the mfr., de-linking and recycling of paper, in
 CC the textile industry, to aid delignification processes, to alter gel
 CC mechanical strength, etc.
 CC

XX Sequence 227 AA;

Query Match 98.4%; Score 1178; DB 17; Length 227;
 Best Local Similarity 100.0%; Pred. No. 8.3e-113;
 Matches 220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KXSVQSAFAATFYGGKDSCTMGACGYGNLYNAGYGLYNALSSALFNDGAMGCACYTI 60
 DB 1 KXSVQSAFAATFYGGKDSCTMGACGYGNLYNAGYGLYNALSSALFNDGAMGCACYTI 60
 OY 61 TCDTQGTWKCKGSGSITTTATNLCPMMLPNSGSGCNPPLXHFDMSPAMENIAYQ 120
 DB 61 TCDTQGTWKCKGSGSITTTATNLCPMMLPNSGSGCNPPLXHFDMSPAMENIAYQ 120
 OY 121 AGIVPVNKKRPVXORSRSGIRFAISGHDFELVTYVNVGSGVVAQMSIKGSNTGMMAMSR 180
 DB 121 AGIVPVNKKRPVXORSRSGIRFAISGHDFELVTYVNVGSGVVAQMSIKGSNTGMMAMSR 180
 OY 181 MWGAMQSNAYLAGOSLSFTVOLDDGRKVTAMNXPXNML 220
 DB 181 MWGAMQSNAYLAGOSLSFTVOLDDGRKVTAMNXPXNML 220

RESULT 2

AAR94527
 ID AAR94527 standard; Protein: 227 AA.

AC AAR94527;

DT 08-JUL-1996 (first entry)

DE Cucumber expansin-29.

XX Expansin-29; plant cell wall; cellulose; paper recycling; de-linking;
 XX polysaccharide; cucumber.
 XX

OS Cucumis sativus var. Burpee Pickler.
 XX

PN A09540262-A.
 XX

PD 04-APR-1996.
 XX

PE 12-MAY-1994; 94AU-0068320.
 XX

PR 12-MAY-1995; 95US-0440517.
 PR 12-MAY-1993; 93US-0060944.

PA (PENN-) PENN STATE RES FOUND.

PI Cosgrove DJ, McQueen-Mason S;

DR WPI: 1996-201150/21.
 DR N-PSDB; AAT13320.
 XX

PT Expansin proteins which alter the mechanical strength of
 PT poly:saccharide(s) - useful in paper mfr. and recycling
 XX
 XX Claim 7; Page 30; 60pp; English.
 PS

CC Cucumber expansin-29 (AAR94527) is a member of a novel class of
 CC proteins that catalyse the extension of plant cell walls and the
 CC weakening of the hydrogen bonds in pure cellulose. It can be obtd.
 CC by expression of an isolated cDNA clone (see AAT13320) in bacterial or
 CC other host cells. Expansin proteins have also been identified in oat
 CC coleoptiles, in Arabidopsis (see AAR94530-32) and in rice (AAR94528-29),
 CC and appear to be broadly distributed throughout the plant kingdom.
 CC Expansins can be used e.g. in the mfr., de-linking and recycling of
 CC paper, in the textile industry, to aid delignification processes, to
 CC alter gel mechanical strength, etc.
 CC

XX Sequence 227 AA;

Query Match 64.6%; Score 773.5; DB 17; Length 227;
 Best Local Similarity 64.5%; Pred. No. 2.8e-71;
 Matches 138; Conservative 27; Mismatches 46; Indels 3; Gaps 2;

OY 6 QSAFATFYGGKDSCTMGACGYGNLYNAGYGLYNALSSALFNDGAMGCACYTTCDS 65
 DB 6 QSHATFYGGGDSAGTMGACGYGNLYSGYGTNTVALSTALPNNGLSCACCEMTC-TN 64

OY 66 QTKKCKRGSGSITTTATNLCPMMLPNSGSGCNPPLXHFDMSPAMENIAYQGIYP 125
 DB 65 DPKWCLPG--TIRVTATNFCPPNFPALPNNNGWCNPLQHFDAEPFLQIAYRAGIYP 122

OY 126 VNYKRPVXORSRSGIRFAISGHDFELVTYVNVGSGVVAQMSIKGSNTGMMAMSR 185
 DB 123 VSFRRVPCMKKGGRFTINGHSYFNLYLITNVGAGAGVHSYIKGSRTGQMSRMMGON 182

OY 186 WQSNAYLAGOSLSFTVOLDDGRKVTAMNXPXNML 219
 DB 183 WQSNAYLAGOSLSFTVOLDDGRKVTAMNXPXNML 216

RESULT 3
 AAG26696
 ID AAG26696 standard; Protein: 252 AA.

AC AAG26696;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 31249.

XX Protein identification: signal transduction pathway; metabolic pathway;
 XX hybridisation assay; genetic mapping; gene expression control; promoter;
 XX termination sequence.
 XX

OS Arabidopsis thaliana.
 XX

PN EPI033405-A2.
 XX

PD 06-SEP-2000.
 XX

PE 25-FEB-2000; 2000EP-0301439.
 XX

PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 11, 2002, 14:49:55 ; Search time 25.1292 Seconds
(without alignments)
1003.367 Million cell updates/sec

Title: US-09-896-301-4

Perfect score: 1197
Sequence: 1 KXSVASAPAFYFGKDGSC.....KVTANNAEPXNMLXXXXXX 227

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_032802.*
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	1178	98.4	227 17 AAR94530	Arabidopsis expans
2	773.5	64.6	227 17 AAR94527	Cucumber expansin
3	766	64.0	252 21 AAG26696	Arabidopsis thalia
4	766	64.0	266 21 AAG26695	Arabidopsis thalia
5	765	63.9	257 22 AAE00413	Tomato seed expans
6	744.5	62.2	241 21 AAG06546	Arabidopsis thalia
7	744.5	62.2	241 21 AAG51633	Arabidopsis thalia
8	744.5	62.2	241 21 AAG51647	Arabidopsis thalia
9	744.5	62.2	249 21 AAG06545	Arabidopsis thalia
10	744.5	62.2	249 21 AAG51632	Arabidopsis thalia
11	744.5	62.2	249 21 AAG51646	Arabidopsis thalia

12	744.5	62.2	259 21 AAG06544	Arabidopsis thalia
13	744.5	62.2	259 21 AAG51631	Arabidopsis thalia
14	744.5	62.2	280 21 AAG51645	Arabidopsis thalia
15	742	62.0	227 17 AAR94532	Arabidopsis expans
16	740.5	61.9	250 22 AAE00414	Tomato seed expans
17	733.5	61.3	228 17 AAR94528	Rice expansin. Or
18	732.5	61.2	257 21 AAG36445	Arabidopsis thalia
19	731.5	61.1	251 21 AAG23852	Arabidopsis thalia
20	731.5	61.1	251 21 AAG43343	Arabidopsis thalia
21	731.5	61.1	253 21 AAG23851	Arabidopsis thalia
22	731.5	61.1	253 21 AAG43342	Arabidopsis thalia
23	731.5	61.1	282 21 AAG43341	Arabidopsis thalia
24	731.5	61.1	282 21 AAG23850	Arabidopsis thalia
25	727.5	60.8	262 21 AAG29931	Arabidopsis thalia
26	727.5	60.8	273 21 AAG29930	Arabidopsis thalia
27	725.5	60.6	255 21 AAG30325	Arabidopsis thalia
28	725.5	60.6	257 21 AAG30324	Arabidopsis thalia
29	719.5	60.1	263 22 AAE00412	Tomato seed expans
30	718	60.0	210 21 AAG26697	Arabidopsis thalia
31	702.5	58.7	250 21 AAG09622	Arabidopsis thalia
32	702.5	58.7	258 21 AAG09621	Arabidopsis thalia
33	702.5	58.7	280 21 AAG09620	Arabidopsis thalia
34	701.5	58.6	241 21 AAG05453	Arabidopsis thalia
35	701.5	58.6	249 21 AAG05452	Arabidopsis thalia
36	701.5	58.6	255 21 AAG05451	Arabidopsis thalia
37	701	58.6	253 21 AAG25443	Arabidopsis thalia
38	701	58.6	253 21 AAG25443	Arabidopsis thalia
39	700.5	58.5	242 21 AAG36570	Arabidopsis thalia
40	700.5	58.5	249 21 AAG36569	Arabidopsis thalia
41	700.5	58.5	255 21 AAG36568	Arabidopsis thalia
42	678	56.6	185 21 AAG27400	Arabidopsis thalia
43	678	56.6	191 21 AAG27399	Arabidopsis thalia
44	677.5	56.6	210 21 AAG36446	Arabidopsis thalia
45	673.5	56.3	210 21 AAG30326	Arabidopsis thalia

ALIGNMENTS

RESULT 1	
AAR94530	standard; Protein; 227 AA.
XX	
XX	
AC	AAR94530:
XX	
DT	08-JUL-1996 (first entry)
XX	
DE	Arabidopsis expansin.
XX	
KW	Expansin; plant cell wall; cellulose; paper recycling; de-linking;
KW	polysaccharide.
XX	
OS	Arabidopsis sp.
XX	
FH	Key
FT	Misc-difference 2
FT	/note= "unidentified amino acid"
FT	Misc-difference 86
FT	/note= "unidentified amino acid"
FT	Misc-difference 104
FT	/note= "unidentified amino acid"
FT	Misc-difference 133
FT	/note= "unidentified amino acid"
FT	Misc-difference 214
FT	/note= "unidentified amino acid"
FT	Misc-difference 217
FT	/note= "unidentified amino acid"
FT	Misc-difference 221..227
FT	/note= "unidentified amino acids"
XX	
XX	
PN	AU9540262-A.
XX	
PD	04-APR-1996.

l

k

RESULT 14

SOFTWARE: PatentIn Ver. 2.1
SEO_ID NO 6

APPLICANT: Gullinan, Mark J
APPLICANT: Shcherban, Tatyana
TITLE OF INVENTION: PURIFIED EXPANSIN PROTEINS
FILE REFERENCE: 1194/1C11A053
CURRENT APPLICATION NUMBER: US/09/092.160C
CURRENT FILING DATE: 1998-06-05
EARLIER APPLICATION NUMBER: 08/440,517
EARLIER FILING DATE: 1995-05-12
EARLIER APPLICATION NUMBER: 08/242,090
EARLIER FILING DATE: 1994-05-12
EARLIER APPLICATION NUMBER: 08/060,944
EARLIER FILING DATE: 1993-05-12
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 225
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Arabidopsis
US-09-092-160-5

Query Match
Best Local Similarity 57.8%; Score 660.5; DB 4; Length 225;
Matches 122; Conservative 33; Mismatches 49; Indels 7; Gaps 4;

QY 10 ATEYGGKSGCTMGACGAGYGLYNAGYGLYNALSSALFNDGAMGACATTTCTDTSQTKW 69
DB 10 ATEYGGADAGTGMGACGAGYGLYNAGYGLYNALSSALFNDGAMGACATTTCTDTSQTKW 68

QY 70 CKPGNSITTTATNLCXPNMALPNSGWCNPPPLXHFDMOSOPAMENIAYVQAGIVPNYK 129
DB 69 CIPGS--IIVRYNLA--NEALANDNGWCNPPPLXHFDMOEPALQIAYRAGIVPAFR 123

QY 130 RVPKRSRGIRFAISGDYELVTVTWGSGVVAQMSIKSGNTG--WMASRMGMAMQS 188
DB 124 RVPCKRGKGIKFTINGRPYDVLITVWAGAGIDRAVSLKSGKTDWQMSRMGMAMQS 183

QY 189 NAVLAGOSLSFIVQLDDRRVTAWNAKXNM 219
DB 184 NTVLRGQSLSFQVYTDSDGRVTVSYDVYPHDM 214

RESULT 11
US-08-845-539-2
Sequence 2, Application US/08845539
Patent No. 5929303
GENERAL INFORMATION:
APPLICANT: Bennett, Alan B.
APPLICANT: Rose, Jocelyn K.C.
TITLE OF INVENTION: Fruit-Specific and Ripening-Regulation
TITLE OF INVENTION: Expansin Genes to Control Fruit Texture and Softening
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/845,539
FILING DATE: 25-APR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Bastian, Kevin L.

REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 023070-078200US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 225 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-845-539-2

Query Match
Best Local Similarity 59.4%; Score 658.5; DB 2; Length 225;
Matches 114; Conservative 32; Mismatches 45; Indels 1; Gaps 1;

QY 6 QSAFATFYGGKSGCTMGACGAGYGLYNAGYGLYNALSSALFNDGAMGACATTTCTDTSQTKW 64
DB 34 ETNATFYGGSDASGTMGACGAGYGLYNAGYGLYNALSSALFNDGAMGACATTTCTDTSQTKW 93

QY 65 SOTKCKPGNSITTTATNLCXPNMALPNSGWCNPPPLXHFDMOSOPAMENIAYVQAGIV 124
DB 94 PNWKWCLPNSPILITATNFCPPNYALPNDNGWCNPPRPHFDLAMPFLKLAQYRAGIV 153

QY 125 PVYKRVKVPKRSRGIRFAISGDYELVTVTWGSGVVAQMSIKSGNTGMMAMSRMGMCA 184
DB 154 PVYKRVKVPKRSRGIRFAISGDYELVTVTWGSGVVAQMSIKSGNTGMMAMSRMGMCA 213

QY 185 WMQSNNAVLAGOS 196
DB 214 WMQSNNAVLAGOS 225

RESULT 12
US-08-440-517A-3
Sequence 3, Application US/08440517A
Patent No. 5959082
GENERAL INFORMATION:
APPLICANT: COSGROVE, DANIEL J.;
APPLICANT: GULLINAN, MARK;
APPLICANT: SHCHERBAN, TATYANA;
APPLICANT: SHI, JUN
TITLE OF INVENTION: PURIFIED EXPANSIN PROTEINS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: INTELLECTUAL PROPERTY OFFICE, THE
ADDRESS: PENNSYLVANIA STATE UNIVERSITY
STREET: 113 TECHNOLOGY CENTER
CITY: UNIVERSITY PARK
STATE: PENNSYLVANIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 16802-7000
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: NEC 286
OPERATING SYSTEM: DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/440,517A
FILING DATE:
CLASSIFICATION: 530
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 222
TYPE: AMINO ACID
TOPOLOGY: UNKNOWN
US-08-440-517A-3

Query Match
Best Local Similarity 54.2%; Score 649; DB 2; Length 222;
Matches 123; Conservative 25; Mismatches 55; Indels 10; Gaps 3;

QY 181 MWGANMOSNAYLAGOSLSFTIYOLDGRRKVTAMNXPXNML 220
Db 181 MWGANMOSNAYLAGOSLSFTIYOLDGRRKVTAMNXPXNML 220

RESULT 2

US-09-092-160-4
Sequence 4, Application US/09092160C
Patent No. 6255466
GENERAL INFORMATION:
APPLICANT: Cosgrove, Daniel J
APPLICANT: McQueen-Mason, Simon
APPLICANT: Gullitman, Mark J
APPLICANT: Shcherban, Tatyana
TITLE OF INVENTION: PURIFIED EXPANSIN PROTEINS
FILE REFERENCE: 1194/1C114US3
CURRENT APPLICATION NUMBER: US/09/092,160C
CURRENT FILING DATE: 1998-06-05
EARLIER APPLICATION NUMBER: 08/440,517
EARLIER FILING DATE: 1995-05-12
EARLIER APPLICATION NUMBER: 08/242,090
EARLIER FILING DATE: 1994-05-12
EARLIER APPLICATION NUMBER: 08/060,944
EARLIER FILING DATE: 1993-05-12
NUMBER OF SEQ ID NOS: 7
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 4
LENGTH: 227
TYPE: PRF
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Arabidopsis
FEATURE:
OTHER INFORMATION: expansin
NAME/KEY: UNSURE
LOCATION: (2)..(227)
OTHER INFORMATION: Xaa is unknown or other.
US-09-092-160-4

Query Match 98.4%: Score 1178; DB 4; Length 227;
Best Local Similarity 100.0%: Pred No. 3.2e-107;
Matches 220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KXSVAGSAFATFYGGKDGSCCTMGACGYNLVYAGVGLYNAALSSALFNDGAMGACVTTI 60
Db 1 KXSVAGSAFATFYGGKDGSCCTMGACGYNLVYAGVGLYNAALSSALFNDGAMGACVTTI 60
QY 61 TCDSOTKCKPGNSITITATNLCXPNMALPNSGGMCPPLXHPMSQPAVENTAVYQ 120
Db 61 TCDSOTKCKPGNSITITATNLCXPNMALPNSGGMCPPLXHPMSQPAVENTAVYQ 120
QY 121 AGIVPVYKRRVXORSIGIRFALISGHDYFELVTVTNVGGSGVVAQMSIKSNTGMMAMSR 180
Db 121 AGIVPVYKRRVXORSIGIRFALISGHDYFELVTVTNVGGSGVVAQMSIKSNTGMMAMSR 180
QY 181 MWGANMOSNAYLAGOSLSFTIYOLDGRRKVTAMNXPXNML 220
Db 181 MWGANMOSNAYLAGOSLSFTIYOLDGRRKVTAMNXPXNML 220

RESULT 3

US-09-092-160-7
Sequence 7, Application US/09092160C
Patent No. 6255466
GENERAL INFORMATION:
APPLICANT: Cosgrove, Daniel J
APPLICANT: McQueen-Mason, Simon
APPLICANT: Gullitman, Mark J
APPLICANT: Shcherban, Tatyana
TITLE OF INVENTION: PURIFIED EXPANSIN PROTEINS

FILE REFERENCE: 1194/1C114US3
CURRENT APPLICATION NUMBER: US/09/092,160C
CURRENT FILING DATE: 1998-06-05
EARLIER APPLICATION NUMBER: 08/440,517
EARLIER FILING DATE: 1995-05-12
EARLIER APPLICATION NUMBER: 08/242,090
EARLIER FILING DATE: 1994-05-12
EARLIER APPLICATION NUMBER: 08/060,944
EARLIER FILING DATE: 1993-05-12
NUMBER OF SEQ ID NOS: 7
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 7
LENGTH: 227
TYPE: PRF
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: cucurbit
US-09-092-160-7

Query Match 64.1%: Score 767.5; DB 4; Length 227;
Best Local Similarity 64.0%: Pred No. 2.9e-67;
Matches 137; Conservative 27; Mismatches 47; Indels 3; Gaps 2;

QY 6 QSAFATFYGGKDGSCCTMGACGYNLVYAGVGLYNAALSSALFNDGAMGACVTTI 65
Db 6 QSAFATFYGGKDGSCCTMGACGYNLVYAGVGLYNAALSSALFNDGAMGACVTTI 65
QY 66 QPKWCKPGNSITITATNLCXPNMALPNSGGMCPPLXHPMSQPAVENTAVYQ 125
Db 66 QPKWCKPGNSITITATNLCXPNMALPNSGGMCPPLXHPMSQPAVENTAVYQ 125
QY 126 VNKRRVXORSIGIRFALISGHDYFELVTVTNVGGSGVVAQMSIKSNTGMMAMSR 185
Db 126 VNKRRVXORSIGIRFALISGHDYFELVTVTNVGGSGVVAQMSIKSNTGMMAMSR 185
QY 186 WOSNAYLAGOSLSFTIYOLDGRRKVTAMNXPXNML 219
Db 186 WOSNAYLAGOSLSFTIYOLDGRRKVTAMNXPXNML 219

US-08-440-517A-6
Sequence 6, Application US/08440517A
Patent No. 5959082
GENERAL INFORMATION:
APPLICANT: COSGROVE, DANIEL J.,
APPLICANT: GULLITMAN, MARK,
APPLICANT: SCHERBAN, TATYANA;
APPLICANT: SHI, JUN
TITLE OF INVENTION: PURIFIED EXPANSIN PROTEINS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: INTELLECTUAL PROPERTY OFFICE, THE
STREET: 113 TECHNOLOGY CENTER
CITY: UNIVERSITY PARK
STATE: PENNSYLVANIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 16802-7000
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: NEC 286
OPERATING SYSTEM: DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/440,517A
FILING DATE:
CLASSIFICATION: 530
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 226
TYPE: AMINO ACID

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 11, 2002, 14:54:05 ; Search time 8.87897 Seconds
(without alignments)
624.466 Million cell updates/sec

Title: US-09-896-301-4

Sequence: 1 KXSVAQSAFATFYGGKDSC.....KVTAWNXAPXNWLXXXXXXXX 227

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

Database : Issued_Patents_AA:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1178	98.4	227	2	US-08-440-517A-4	Sequence 4, Appl1
2	1178	98.4	227	4	US-09-092-160-7	Sequence 4, Appl1
3	767.5	64.1	227	4	US-09-092-160-7	Sequence 7, Appl1
4	742	62.0	226	2	US-08-440-517A-6	Sequence 6, Appl1
5	742	62.0	226	4	US-09-092-160-6	Sequence 6, Appl1
6	733.5	61.3	228	2	US-08-440-517A-2	Sequence 2, Appl1
7	733.5	61.3	228	4	US-09-092-160-2	Sequence 2, Appl1
8	663.5	55.4	225	4	US-09-362-642-2	Sequence 2, Appl1
9	660.5	55.2	225	2	US-08-440-517A-5	Sequence 5, Appl1
10	660.5	55.2	225	4	US-09-092-160-5	Sequence 5, Appl1
11	658.5	55.0	225	2	US-08-845-539-2	Sequence 2, Appl1
12	649	54.2	222	2	US-08-440-517A-3	Sequence 3, Appl1
13	649	54.2	222	4	US-09-092-160-3	Sequence 3, Appl1
14	627.5	52.4	179	2	US-08-845-539-6	Sequence 6, Appl1
15	627.5	52.4	179	4	US-09-362-642-6	Sequence 6, Appl1
16	576.5	48.2	167	2	US-08-845-539-4	Sequence 4, Appl1
17	576.5	48.2	167	4	US-09-362-642-4	Sequence 4, Appl1
18	180.5	15.1	261	1	US-07-971-096-2	Sequence 2, Appl1
19	180.5	15.1	261	1	US-08-175-096-2	Sequence 2, Appl1
20	169.5	14.2	263	1	US-07-971-096-4	Sequence 4, Appl1
21	169.5	14.2	263	1	US-08-175-096-4	Sequence 4, Appl1
22	169.5	14.2	263	4	US-08-413-974-6	Sequence 6, Appl1
23	169.5	14.2	263	4	US-08-434-418-6	Sequence 6, Appl1
24	169.5	14.2	263	4	US-08-433-388-6	Sequence 6, Appl1
25	169.5	14.2	263	4	US-08-174-739A-6	Sequence 6, Appl1
26	165	13.8	246	4	US-08-441-507-21	Sequence 11, Appl1
27	165	13.8	272	4	US-08-441-507-15	Sequence 15, Appl1

ALIGNMENTS

28	160	13.4	225	4	US-08-441-507-24	Sequence 24, Appl
29	125.5	10.5	197	4	US-08-441-507-5	Sequence 5, Appl1
30	125.5	10.5	200	4	US-08-441-507-4	Sequence 4, Appl1
31	107.5	9.0	122	4	US-08-441-507-23	Sequence 23, Appl1
32	92.5	7.7	213	5	US-08-441-507-31	Sequence 31, Appl1
33	92.5	7.7	123	5	US-08-441-507-31	Sequence 31, Appl1
34	92.5	7.7	134	1	US-08-486-036A-4	Sequence 4, Appl1
35	92.5	7.7	134	5	US-08-486-036A-4	Sequence 4, Appl1
36	92.5	7.7	356	1	US-08-486-036A-2	Sequence 2, Appl1
37	92.5	7.7	356	1	US-08-486-036A-2	Sequence 2, Appl1
38	90.5	7.6	316	1	US-08-482-282B-2	Sequence 2, Appl1
39	90.5	7.6	400	1	US-08-482-282B-2	Sequence 2, Appl1
40	90.5	7.0	145	4	US-08-413-97A-4	Sequence 4, Appl1
41	83.5	7.0	145	4	US-08-434-118-4	Sequence 4, Appl1
42	83.5	7.0	145	4	US-08-434-288-4	Sequence 4, Appl1
43	83.5	7.0	145	4	US-08-174-735A-4	Sequence 4, Appl1
44	82.5	6.9	284	1	US-08-411-777-10	Sequence 10, Appl1
45	82.5	6.9	284	3	US-09-057-088-10	Sequence 10, Appl1

RESULT 1
US-08-440-517A-4

; Sequence 4, Application US/08440517A
; Patent No. 5959082

APPLICANT: COSGROVE, DANIEL J.

APPLICANT: GUILTINAN, MARK;
APPLICANT: SCHERBAN, TATYANA;

Query Match	98.4%	Score 1178	DB 2	Length 227
Best Local Similarity	100.0%	Pred. No. 3.2e-107		
Matches 220; Conservative	0	Mismatches 0	Indels 0	Gaps 0

Qy	1	KXSVASQSAFAAFYFGGKDCSCFMGACGCGNLYNMGYGLYNALSSALFENGGAMGACYYI	60
Db	1	KXSVASQSAFAAFYFGGKDCSCFMGACGCGNLYNMGYGLYNALSSALFENGGAMGACYYI	60
Qy	1	KXSVASQSAFAAFYFGGKDCSCFMGACGCGNLYNMGYGLYNALSSALFENGGAMGACYYI	60
Db	1	KXSVASQSAFAAFYFGGKDCSCFMGACGCGNLYNMGYGLYNALSSALFENGGAMGACYYI	60
Qy	61	TCDDTSQRTWCKPCPGNSTITTTATNLCCPMMALPSSMSGCNCPLKHPDMSAPAMENINAYQ	120
Db	61	TCDDTSQRTWCKPCPGNSTITTTATNLCCPMMALPSSMSGCNCPLKHPDMSAPAMENINAYQ	120
Qy	121	AGIVPVNKKRPVXPXORSGIRFPAISGHDYFELVPTVNNVGGSGVAQOMSTKGSNTGWMAMSR	180
Db	121	AGIVPVNKKRPVXPXORSGIRFPAISGHDYFELVPTVNNVGGSGVAQOMSTKGSNTGWMAMSR	180

QY 185 WMSNAYLAGOSLSFTVOLDGGRKVTAMNXPXNM 219
 Db 217 WMSNAYLAGOSLSFTVOLDGGRKVTAMNXPXNM 251

RESULT 13

expansin 18 - tomato
 C:Species: Lycopersicon esculentum (tomato)
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 20-Jun-2000
 C:Accession: T06573
 R:Reinhardt, D.; Wiltner, F.; Mandel, T.; Kuhlmeier, C.
 submitted to the EMBL Data Library, March 1998
 A:Description: Localized up-regulation of a new expansin gene predicts the site of leaf
 A:Reference number: 215768
 A:Accession: T06573
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-255 <RE>
 A:Cross-references: EMBL:AJ004997; PIDN:CAA06271.1
 A:Experimental source: cultivar Money Maker
 C:Genetics:
 A:Gene: exp18
 C:Superfamily: expansin

Query Match 61.7%; Score 738.5; DB 2; Length 255;
 Best Local Similarity 61.7%; Pred. No. 2,7e-56;
 Matches 132; Conservative 30; Mismatches 51; Indels 1; Gaps 1;

QY 6 QSAFAFFYGKDGSCITMGACGYGLNAGYGLYNALSSALFNDGAMCGACTTTCDS 65
 Db 30 QSAHAFFYGGSDASGTMGACGYGLNAGYGLYNALSSALFNDGAMCGACTTTCDS 88
 QY 66 QTKWCKPGGNSITTTATNLCXRNMAIPSSGCGWCPPLXHPDMSOPAMENIAVYQAGIYP 125
 Db 89 YPQWCHPGSPSIFITATNCPNFPALPNDGCGWCPPLXHPDMSOPAMENIAVYQAGIYP 148
 QY 126 VVKRVPXORSRGIRFPAISGHDYFELVTVTNVGGSGVVAQMSIKGSNTGMMAMSRNMGAN 185
 Db 149 VYRRPCKRKGIRFTINGFRFNLVLTNVAAGDILSLIKSKTWMISNRNMGON 208
 QY 186 WOSNAYLAGOSLSFTVOLDGGRKVTAMNXPXNM 219
 Db 209 WQTNVYLFQGSLSFRVRADRRSSTSMITPAHW 242

RESULT 14

expansin (clone pTexp4) - loblolly pine (fragment)
 C:Species: Pinus taeda (loblolly pine)
 C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 26-May-2000
 C:Accession: T09825
 R:Hutchison, K.W.; Singer, P.B.; Diaz-Sala, C.; Greenwood, M.S.
 submitted to the EMBL Data Library, July 1996
 A:Description: Expansins are conserved in conifers and expressed in response to exogenous
 A:Reference number: 216866
 A:Accession: T09825
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-232 <HUT>
 A:Cross-references: EMBL:U64892; NID:91778102; PID:91778103
 A:Experimental source: clone pTexp4
 C:Superfamily: expansin

Query Match 61.6%; Score 737.5; DB 2; Length 232;
 Best Local Similarity 63.1%; Pred. No. 3e-56;
 Matches 135; Conservative 24; Mismatches 52; Indels 3; Gaps 2;

QY 6 QSAFAFFYGKDGSCITMGACGYGLNAGYGLYNALSSALFNDGAMCGACTTTCDS 65
 Db 11 QSAHAFFYGGSDASGTMGACGYGLNAGYGLYNALSSALFNDGAMCGACTTTCDS 69
 QY 66 QTKWCKPGGNSITTTATNLCXRNMAIPSSGCGWCPPLXHPDMSOPAMENIAVYQAGIYP 125

Db 70 DPQWCLPG--VTITATNCPNFPALPNDGCGWCPPLXHPDMSOPAMENIAVYQAGIYP 127
 QY 126 VVKRVPXORSRGIRFPAISGHDYFELVTVTNVGGSGVVAQMSIKGSNTGMMAMSRNMGAN 185
 Db 128 ILYTRVPCIRKGIREFTVNGHSCFNLVLTNVAAGDILSLIKSKTWMISNRNMGON 187
 QY 186 WOSNAYLAGOSLSFTVOLDGGRKVTAMNXPXNM 219
 Db 188 WOSNAYLAGOSLSFTVOLDGGRKVTAMNXPXNM 221

RESULT 15

expansin S2 precursor - cucumber
 C:Species: Cucumis sativus (cucumber)
 C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
 C:Accession: T10083
 R:Shcherban, T.Y.; Shi, J.; Durachko, D.M.; Gullitlan, M.J.; McQueen-Mason, S.J.; Shi
 Proc. Natl. Acad. Sci. U.S.A. 92, 9245-9249, 1995
 A:Title: Molecular cloning and sequence analysis of expansins--a highly conserved, mu
 A:Reference number: 214894; MUID:96016146
 A:Accession: T10083
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-258 <SHC>
 A:Cross-references: EMBL:U030460; NID:91040876; PIDN:AA837749.1; PID:91040877
 A:Experimental source: cv. Burpee Pickler; hypocotyl
 C:Genetics:
 A:Gene: EXP2
 C:Function:
 A:Description: induces cell wall extension in plants
 C:Superfamily: expansin
 F:1-24/Domain: signal sequence #status predicted <SIG>
 F:25-258/Product: expansin #status predicted <MAT>

Query Match 61.5%; Score 736.5; DB 2; Length 258;
 Best Local Similarity 61.7%; Pred. No. 4.1e-56;
 Matches 132; Conservative 28; Mismatches 51; Indels 3; Gaps 3;

QY 8 AFATFYGGKDGSCITMGACGYGLNAGYGLYNALSSALFNDGAMCGACTTTCDS-TSQ 66
 Db 33 AHATFYGGSDASGTMGACGYGLNAGYGLYNALSSALFNDGAMCGACTTTCDS-KTD 92
 QY 67 TKWCKPGGNSITTTATNLCXRNMAIPSSGCGWCPPLXHPDMSOPAMENIAVYQAGIYP 126
 Db 93 PRWCIRKA-SVTITATNCPNFPALPNDGCGWCPPLXHPDMSOPAMENIAVYQAGIYP 151
 QY 127 VVKRVPXORSRGIRFPAISGHDYFELVTVTNVGGSGVVAQMSIKGS-NTGMMAMSRNMGAN 185
 Db 152 LYQRPCKRKGIRFTINGFRFNLVLTNVAAGDILSLIKSKTWMISNRNMGON 211
 QY 186 WOSNAYLAGOSLSFTVOLDGGRKVTAMNXPXNM 219
 Db 212 WOSNAYLAGOSLSFTVOLDGGRKVTAMNXPXNM 245

Search completed: October 11, 2002, 15:01:00
 Job time : 12.3919 secs

	Matches	137;	Conservative	21;	Mismatches	54;	Indels	0;	Gaps	0;	
QY	8	AFAATFFGKDDSC	CTMGACGCTG	NTLNAGTGLYNAL	SSALFPNDGAM	CGACYTIT	GGDTSOT	67			
Db	37	AFAATFFGSDAS	GTMGACGYG	DTYSTGYTAL	TVLEFNDGAS	CGCYTRIMCD	YQAD	96			
QY	68	KWCKPFGNSI	ITITATNL	CXPMNALP	SNSGMGN	CPPLXHF	DMSPAPAMEN	IAVYQAGI	VPVN	127	
Db	97	RRCFSGISG	VTITATNL	CPNPALE	PNDGAGM	CNPNPRQ	HFDMAE	PAEMLIK	IGVYVG	GI	156
QY	128	YKRVFXORS	GITREAI	ISCHDYF	ELVTVTN	VGSGVYA	QMSIKGS	NTGMMAMS	RNNG	ANNO	187
Db	157	YQVBCAQQG	GVREFIT	INGRDYF	ELVTVTN	VGSGVSI	QMSIKGS	RTGMMAMS	RNNG	VMMQ	216
QY	188	SNAYLAGOS	LSPIVGL	DDGRK	TVANXXA	PXNN	219				
Db	217	SNAYLAGOS	LSKFTV	SSDGOTIT	FLDVAP	ACH	248				

```

RESULT 6
T09821
expansin (clone ptxexp3) - loblolly pine (fragment)
C:Species: Pinus taeda (loblolly pine)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 26-May-2000
C:Accession: T09821
R:Hutchinson, K.W.; Singer, P.B.; Diaz-Sala, C.; Greenwood, M.S.
submitted to the EMBL Data Library, July 1996
A:Description: Expansins are conserved in conifers and expressed in response to exogenous
A:Reference number: Z16866
A:Accession: T09821
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-232 <HUP>
A:Cross-references: EMBL:U64891; NTD:g1778100; PTD:g1778101
A:Experimental source: clone ptxexp3
A:Superfamily: expansin

```

[illegible]

RESULT 7
T09826
expansin (clone pTexp5) - loblolly pine (fragment)
C:Species: Pinus taeda (loblolly pine)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 26-May-2000
C:Accession: T09826
R:Hutchinson, K.W.; Slinger, P.B.; Diaz-Sala, C.; Greenwood, M.S.
submitted to the EMBL Data Library July 1996
A:Description: Expansins are conserved in conifers and expressed in response to exogenous
A:Reference number: Z16866
A:Accession: T09826
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-232 <HUT>
A:Cross-references:EMBL:U64893; NID:g1778104; PID:g1778105

A;Experimental source: clone pTexp5
C;Superfamily: expansin

Query Match	62.7%	Score	750.5	DB	2	Length	232
Best Local Similarity	63.6%	Pred.	No. 2.3e-57				
Matches	136	Conservative	25	Mismatches	50	Indels	3
						Gaps	2

QY	6	OSAFETFFGKGDSCGTMGACGYNINAGLGNALNSALNFEDGSMGACVYITDTS	63
QY	6	OSAFETFFGKGDSCGTMGACGYNINAGLGNALNSALNFEDGSMGACVYITDTS	63
Db	11	ESAAHTFFGKGDSCGTMGACGYNINAGLGNALNSALNFEDGSMGACVYITDTS	69
QY	66	QTKMCKPGNSITTTATNLCKPNNALPSNSGWCNCPPLXHEFMSQPAEMINAYQAGIVP	125
Db	70	DPQWCLPG-FTYVATATFCPPNNALPFDNGMGNPNPLOHEDMAEPALFIKARGGIVP	127
QY	126	VNYKRVPKQBSGRIEALSGHDYELVYTVNWGSSGVYAQMSIKGNTGMWMSRRNGAN	185
QY	128	ILYTVPECLRKRGIFETVNGHSTFNLVYTVNWGAGADVAIVSLKSGSGWQPMPSRRNGON	187
QY	186	WQSNAYLAGOSLSFTVQDDSRKTYAMNAXPXN	219
Db	188	WQSNAYLAGOSLSFTVQDDSRKTYAMNAXPXN	221

RESULT 8
 T09818
 expansin (clone ptxexp2) - loblolly pine (fragment)
 C:Species: Pinus taeda (loblolly pine)
 C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 26-May-2000
 C:Accession: T09818
 R:Hutchinson, K.W.; Singer, P.B.; Diaz-Sala, C.; Greenwood, M.S.
 submitted to the EMBL Data Library, July 1996
 A:Description: Expansins are conserved in conifers and expressed in response to exogen
 A:Reference number: Z16866
 A:Accession: T09818
 A:Status: translated from GB/EMBL/DBDU
 A:Molecule type: mRNA
 A:Residues: 1-232 <NR>
 A:Cross-references: EMBL:U64890; NID:g1778098; PID:g1778099
 A:Experimental source: clone ptxexp2; hypocotyl
 C:Superfamily: expansin

[illegible]

```

RESULT 9
T50658
expansin 9 [imported] - tomato
C:Species: lycopersicon esculentum (tomato)
C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #ext_change 28-Jul-2000
C:Accession: T50658
R:Caderras, D.M.
submitted to the EMBL Data Library, June 1999
A:Reference number: Z25160
A:Accession: T50658

```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 11, 2002, 14:53:05 ; Search time 11.3919 Seconds
(without alignments)
1914.721 Million cell updates/sec

Title: US-09-896-301-4

Perfect score: 1197
Sequence: 1 KXSAQSAFAATFYGGKDGSC.....KVTAMNAPXNMLXXXXXX 227

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: PIR_71:*
2: PIR1:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1173	98.0	255	T03299	expansin 3 - rice
2	773.5	64.6	250	T10079	expansin S1 precu
3	764.5	63.9	252	F86335	hypothetical prote
4	762.5	63.7	251	T03298	expansin 2 - rice
5	756	63.2	261	T03737	expansin - rice
6	750.5	62.7	232	T09821	expansin (clone pp
7	745.5	62.7	232	T09826	expansin (clone pp
8	747.5	62.4	232	T09818	expansin (clone pp
9	746.5	62.4	257	T50658	expansin 9 (import
10	746.5	62.4	260	T47689	expansin 1-like prot
11	743.5	62.1	246	T04175	expansin - rice
12	742.5	62.0	264	T50659	alpha-expansin OSE
13	738.5	61.7	255	T06573	expansin 18 - toma
14	737.5	61.6	232	T09825	expansin (clone pp
15	736.5	61.5	258	T10083	expansin S2 precu
16	733.5	61.3	237	T50634	expansin Expi (imp
17	733.5	61.3	262	T50660	alpha-expansin 2 (
18	732.5	61.2	257	D84820	probable expansin
19	731.5	61.1	248	C84444	pollen allergen ho
20	731.5	61.1	258	S53082	probable expansin
21	727.5	60.8	262	T02530	probable expansin
22	725.5	60.6	257	T02727	expansin EXP6 (imp
23	725.5	60.6	259	T50653	expansin 1 - tomat
24	721.5	60.3	261	T07630	probable expansin
25	709.5	59.3	258	T09786	expansin-like prot
26	707.5	59.1	260	T08016	expansin - upland
27	702.5	58.7	258	T48247	expansin EXP2 (imp
28	702	58.6	253	T50656	probable expansin
29	701	58.6	253	F84831	probable expansin

30	692.5	57.9	255	2	T02010	expansin homolog T
31	670	56.0	255	2	T50655	expansin EXP5 (imp
32	626.5	52.3	257	2	F86259	protein T12C24.10
33	600.5	50.2	257	2	G96654	hypothetical prote
34	513	42.9	160	2	T09871	expansin - upland
35	503.5	42.1	256	2	T05648	expansin homolog F
36	233	19.5	491	2	F96681	protein F1R22.6 (1
37	229.5	19.2	259	2	T50657	beta-expansin (imp
38	229.5	19.2	271	2	H84592	beta-expansin (imp
39	221	18.5	259	2	E84886	probable beta-expa
40	219	18.3	77	2	T09815	expansin (clone pp
41	217	18.1	102	2	T09828	expansin (clone pp
42	213.5	17.8	277	2	S48032	ciml protein - soy
43	207.5	17.3	276	2	T09041	ciml protein homol
44	199	16.6	81	2	T09830	expansin (clone pp
45	197.5	16.5	261	2	T04301	beta-expansin - r1

ALIGNMENTS

RESULT 1
T03299
expansin 3 - rice
C:Species: Oryza sativa (rice)
C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 21-Jul-2000
C:Accession: T03299
R:Shcherban, T.Y.; Shi, J.; Durachko, D.M.; Gultinan, M.J.; McQueen-Mason, S.J.; Shi
Proc. Natl. Acad. Sci. U.S.A. 92, 9245-9249, 1995
A:Title: Molecular cloning and sequence analysis of expansins--a highly conserved, mu
A:Reference number: Z14894; MUID:96016146
A:Accession: T03299
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-255 <SHC>
A:Cross-references: EMBL:U030479; NID:G1041711; PTDN:AAB38075.1; PID:G1041712
C:Genetics:
A:Gene: EXP3
C:Function:
A:Description: induces extension (creep) in plant cell walls
C:Superfamily: expansin

Query Match
Best Local Similarity 98.0%; Score 1173; DB 2; Length 255;
Matches 214; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY	1	KXSAQSAFAATFYGGKDGSCITMGACGYGNDLYNMGYGLYNALSSALFNDGAMCGACTT	60
DB	24	KXSAQSAFAATFYGGKDGSCITMGACGYGNDLYNMGYGLYNALSSALFNDGAMCGACTT	83
QY	61	TCDDTSQRTWCKRPGGSGITTTNTNLCXPRMALPNSGCGNPLXHPMSQPAWENIAYQ	120
DB	84	TCDDTSQRTWCKRPGGSGITTTNTNLCXPRMALPNSGCGNPLXHPMSQPAWENIAYQ	143
QY	121	AGIYVNVKRPVXORSGIRFASGHDFELVTYTNVGSQVVAOMSTKSGNTGMAMSR	180
DB	144	AGIYVNVKRPVXORSGIRFASGHDFELVTYTNVGSQVVAOMSTKSGNTGMAMSR	203
QY	181	NWGANWQSNAYLAGOSLSFIVOLDGKRKVTAMNAPXNM	219
DB	204	NWGANWQSNAYLAGOSLSFIVOLDGKRKVTAMNAPXNM	242

RESULT 2
T10079
expansin S1 precursor - cucumber
C:Species: Cucumis sativus (cucumber)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C:Accession: T10079
R:Shcherban, T.Y.; Shi, J.; Durachko, D.M.; Gultinan, M.J.; McQueen-Mason, S.J.; Shi
Proc. Natl. Acad. Sci. U.S.A. 92, 9245-9249, 1995
A:Title: Molecular cloning and sequence analysis of expansins--a highly conserved, mu
A:Reference number: Z14894; MUID:96016146

RESULT 15

```

ENV2_MOUSE      STANDARD;      PRT;      432 AA.
ID  ENV2_MOUSE
AC  P11370;
DT  01-JUL-1989 (Rel. 11, Created)
DT  01-JUL-1989 (Rel. 11, last sequence update)
DT  01-NOV-1991 (Rel. 20, last annotation update)
DE  Retrovirus-related ENV polypeptide (FV-4 locus) (Fragments).
GN  ENV.
OS  Mus musculus (Mouse).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN  NCBI_TaxID=10090;
RP  [1]
RX  MEDLINE=85265041; PubMed=2991595;
RT  Ikeda H., Laigret F., Martin M.A., Repaske R.;
RT  "Characterization of a molecularly cloned retroviral sequence
RT  associated with FV-4 resistance.";
RT  J. Virol. 55:768-777(1985).
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@sib-sib.ch).
CC  -----
DR  EMBL; M11051; AAA37562.1; -.
DR  EMBL; M11052; AAA37563.1; -.
DR  PIR; A25483; A25483.
DR  HSSP; P03390; IAOI.
DR  Pfam; PF00429; ENV_polypeptide; 1.
KW  Coat protein; Polypeptide.
FT  NON_CONS 348 349
SQ  SEQUENCE 432 AA; 47661 MW; 092D3963EE5771FC CRC64;

```

Query Match

Best Local Similarity 7.1%; Score 84.5; DB 1; Length 432;
Matches 29; Conservative 10; Mismatches 37; Indels 27; Gaps 5;

```

QY  54 CGA-----CYTTCDFRSGTKWCKPGGNSITIT-ATNLCKXPNMALPNSNGWCNPLXHD 107
    ||      | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB  164 CGCPAYTCASWCCETTGRAAKPVTSSWDYITVSNLSSPQAFKACKNNGCNPLVYRF- 222
QY  108 MSQPAMENIAVYQAGIVPNYKRPVXORSG---GIRFAISGHD 147
    : |      : |      : |      : |      : |      : |      : |      : |
DB  223 -TGPG-----KRATSWTGTGHEWGLRLYISGHD 248

```

Search completed: October 11, 2002, 14:57:32
Job time : 7.86347 secs

FT CARBOHYD 430 430 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT NON_TER 534 534
 SQ SEQUENCE 534 AA: 58348 MW: D3A50B10B7F22BB3 CRC64:

Query Match 7.1%; Score 85.5; DB 1; Length 534;
 Best Local Similarity 21.5%; Pred. No. 2.3;
 Matches 34; Conservative 15; Mismatches 44; Indels 65; Gaps 6;

14 GGDGSCCTMGACGYNLY-----NAGCYLYNNAISSALPNDGAMGACYT 59
 126 GPDDGCAVWGCTGTTGTYRPTSSMDYITVKKGYTOGIQC-----SGGCMGCPYD 178
 60 ITCDTSQTKWCKPGGNSITTTATNLCXPMNALPNSGCMCPPLXHPDMS--OPAWENIA 117
 179 -----KAVHSST-----GASEGRCNPLLIQFOKGRQTSWDG-- 212
 118 VYQAGIVPVNKRVPXORSGRIFRAISGHDFELVTVT 155
 213 -----PKSWGRLVRSGYDPLALFSVS 234

RESULT 13

ENV_FLVLB STANDARD; PRT: 662 AA.

ID ENV_FLVLB 085515; 085516; 085517;

AC P1261; 085515; 085516; 085517;

DT 01-JUL-1989 (Rel. 11, Last sequence update)

DT 01-JUL-1989 (Rel. 11, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE ENV POLYPROTEIN precursor (coat polypeptide) [Contains: knob protein

DE GP70; Spike protein P15E] (Fragment).

GN ENV.

OS Feline leukemia virus (strain B/lambda-B1).

OC Viruses; Retroviral viruses; Retroviridae; Gammaretrovirus.

NCBI_TaxID=103916;

RP SEQUENCE FROM N.A.

RA MEDLINE=8036192; PubMed=2444714;

RA Elider J.H.; Kumar H.P.M., Fitting T., Grant C.K.,

RT "Natural feline leukemia virus variant escapes neutralization by a

RT monoclonal antibody via an amino acid change outside the antibody-

RT binding epitope."

RT J. Virol. 61:3410-3415(1987).

CC -1- PVT: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.

CC -1- PVT: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.

CC -1- PVT: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.

CC -1- PVT: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.

CC -1- PVT: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.

CC -1- PVT: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.

CC -1- PVT: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.

CC -1- PVT: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.

CC -1- PVT: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.

CC -1- PVT: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.

CC -1- PVT: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.

CC -1- PVT: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.

CC -1- PVT: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.

CC -1- PVT: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.

CC -1- PVT: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.

CC -1- PVT: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.

CC -1- PVT: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.

CC -1- PVT: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.

CC -1- PVT: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.

CC -1- PVT: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.

CC -1- PVT: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.

CC -1- PVT: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.

CC -1- PVT: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.

CC -1- PVT: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.

Query Match 7.1%; Score 85.5; DB 1; Length 662;
 Best Local Similarity 21.5%; Pred. No. 2.8;
 Matches 34; Conservative 15; Mismatches 44; Indels 65; Gaps 6;

14 GGDGSCCTMGACGYNLY-----NAGCYLYNNAISSALPNDGAMGACYT 59
 126 GPDDGCAVWGCTGTTGTYRPTSSMDYITVKKGYTOGIQC-----SGGCMGCPYD 178
 60 ITCDTSQTKWCKPGGNSITTTATNLCXPMNALPNSGCMCPPLXHPDMS--OPAWENIA 117
 179 -----KAVHSST-----GASEGRCNPLLIQFOKGRQTSWDG-- 212
 118 VYQAGIVPVNKRVPXORSGRIFRAISGHDFELVTVT 155
 213 -----PKSWGRLVRSGYDPLALFSVS 234

RESULT 14

GUN_MYTED STANDARD; PRT: 181 AA.

ID GUN_MYTED 082186;

AC P82186;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 01-MAR-2002 (Rel. 41, Last annotation update)

DE Endoglucanase (EC 3.2.1.4) (Endo-1,4-beta-glucanase) (Cellulase)

DE (CMCase).

OS Mytilus edulis (Blue mussel).

OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloida;

OC Mytilidae; Mytilidace; Mytilus.

NCBI_TaxID=6530;

RP SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.

RA TISSUE-Digestive gland;

RA MEDLINE=20389517; PubMed=10931178;

RA Xu B., Hellman U., Ersson B., Janson J.-C.;

RT "Purification, characterization and amino acid sequence analysis of a

RT thermostable, low molecular mass endo-beta-1,4-glucanase from blue

RT mussel, Mytilus edulis."

RT Eur. J. Biochem. 267:4970-4977(1999).

CC -1- FUNCTION: ACTIVE TOWARDS THE SOLUBLE CARBOXYMETHYLCELLULOSE (CMC).

CC -1- FUNCTION: ACTIVE TOWARDS THE SOLUBLE CARBOXYMETHYLCELLULOSE (CMC).

CC -1- FUNCTION: ACTIVE TOWARDS THE SOLUBLE CARBOXYMETHYLCELLULOSE (CMC).

CC -1- FUNCTION: ACTIVE TOWARDS THE SOLUBLE CARBOXYMETHYLCELLULOSE (CMC).

CC -1- FUNCTION: ACTIVE TOWARDS THE SOLUBLE CARBOXYMETHYLCELLULOSE (CMC).

CC -1- FUNCTION: ACTIVE TOWARDS THE SOLUBLE CARBOXYMETHYLCELLULOSE (CMC).

CC -1- FUNCTION: ACTIVE TOWARDS THE SOLUBLE CARBOXYMETHYLCELLULOSE (CMC).

CC -1- FUNCTION: ACTIVE TOWARDS THE SOLUBLE CARBOXYMETHYLCELLULOSE (CMC).

CC -1- FUNCTION: ACTIVE TOWARDS THE SOLUBLE CARBOXYMETHYLCELLULOSE (CMC).

CC -1- FUNCTION: ACTIVE TOWARDS THE SOLUBLE CARBOXYMETHYLCELLULOSE (CMC).

CC -1- FUNCTION: ACTIVE TOWARDS THE SOLUBLE CARBOXYMETHYLCELLULOSE (CMC).

CC -1- FUNCTION: ACTIVE TOWARDS THE SOLUBLE CARBOXYMETHYLCELLULOSE (CMC).

CC -1- FUNCTION: ACTIVE TOWARDS THE SOLUBLE CARBOXYMETHYLCELLULOSE (CMC).

CC -1- FUNCTION: ACTIVE TOWARDS THE SOLUBLE CARBOXYMETHYLCELLULOSE (CMC).

CC -1- FUNCTION: ACTIVE TOWARDS THE SOLUBLE CARBOXYMETHYLCELLULOSE (CMC).

CC -1- FUNCTION: ACTIVE TOWARDS THE SOLUBLE CARBOXYMETHYLCELLULOSE (CMC).

CC -1- FUNCTION: ACTIVE TOWARDS THE SOLUBLE CARBOXYMETHYLCELLULOSE (CMC).

CC -1- FUNCTION: ACTIVE TOWARDS THE SOLUBLE CARBOXYMETHYLCELLULOSE (CMC).

CC -1- FUNCTION: ACTIVE TOWARDS THE SOLUBLE CARBOXYMETHYLCELLULOSE (CMC).

CC -1- FUNCTION: ACTIVE TOWARDS THE SOLUBLE CARBOXYMETHYLCELLULOSE (CMC).

CC -1- FUNCTION: ACTIVE TOWARDS THE SOLUBLE CARBOXYMETHYLCELLULOSE (CMC).

CC -1- FUNCTION: ACTIVE TOWARDS THE SOLUBLE CARBOXYMETHYLCELLULOSE (CMC).

CC -1- FUNCTION: ACTIVE TOWARDS THE SOLUBLE CARBOXYMETHYLCELLULOSE (CMC).

CC -1- FUNCTION: ACTIVE TOWARDS THE SOLUBLE CARBOXYMETHYLCELLULOSE (CMC).

CC -1- FUNCTION: ACTIVE TOWARDS THE SOLUBLE CARBOXYMETHYLCELLULOSE (CMC).

CC -1- FUNCTION: ACTIVE TOWARDS THE SOLUBLE CARBOXYMETHYLCELLULOSE (CMC).

Query Match 7.1%; Score 85; DB 1; Length 181;
 Best Local Similarity 32.5%; Pred. No. 0.86;
 Matches 37; Conservative 6; Mismatches 35; Indels 36; Gaps 9;

13 GGDGSCCTMGACGYNLY-----NAGCYLYNNAISSALPNDGAMGACYT 59
 126 GPDDGCAVWGCTGTTGTYRPTSSMDYITVKKGYTOGIQC-----SGGCMGCPYD 178
 60 ITCDTSQTKWCKPGGNSITTTATNLCXPMNALPNSGCMCPPLXHPDMS--OPAWENIA 117
 179 -----KAVHSST-----GASEGRCNPLLIQFOKGRQTSWDG-- 212
 118 VYQAGIVPVNKRVPXORSGRIFRAISGHDFELVTVT 155
 213 -----PKSWGRLVRSGYDPLALFSVS 234

RESULT 15

GUN_MYTED STANDARD; PRT: 181 AA.

ID GUN_MYTED 082186;

AC P82186;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 01-MAR-2002 (Rel. 41, Last annotation update)

DE Endoglucanase (EC 3.2.1.4) (Endo-1,4-beta-glucanase) (Cellulase)

DE (CMCase).

OS Mytilus edulis (Blue mussel).

OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloida;

OC Mytilidae; Mytilidace; Mytilus.

NCBI_TaxID=6530;

RP SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.

RA TISSUE-Digestive gland;

RA MEDLINE=20389517; PubMed=10931178;

RA Xu B., Hellman U., Ersson B., Janson J.-C.;

RT "Purification, characterization and amino acid sequence analysis of a

RT thermostable, low molecular mass endo-beta-1,4-glucanase from blue

RT mussel, Mytilus edulis."

RT Eur. J. Biochem. 267:4970-4977(1999).

CC -1- FUNCTION: ACTIVE TOWARDS THE SOLUBLE CARBOXYMETHYLCELLULOSE (CMC).

CC -1- FUNCTION: ACTIVE TOWARDS THE SOLUBLE CARBOXYMETHYLCELLULOSE (CMC).

CC -1- FUNCTION: ACTIVE TOWARDS THE SOLUBLE CARBOXYMETHYLCELLULOSE (CMC).

CC -1- FUNCTION: ACTIVE TOWARDS THE SOLUBLE CARBOXYMETHYLCELLULOSE (CMC).

CC -1- FUNCTION: ACTIVE TOWARDS THE SOLUBLE CARBOXYMETHYLCELLULOSE (CMC).

CC -1- FUNCTION: ACTIVE TOWARDS THE SOLUBLE CARBOXYMETHYLCELLULOSE (CMC).

CC -1- FUNCTION: ACTIVE TOWARDS THE SOLUBLE CARBOXYMETHYLCELLULOSE (CMC).

CC -1- FUNCTION: ACTIVE TOWARDS THE SOLUBLE CARBOXYMETHYLCELLULOSE (CMC).

CC -1- FUNCTION: ACTIVE TOWARDS THE SOLUBLE CARBOXYMETHYLCELLULOSE (CMC).

CC -1- FUNCTION: ACTIVE TOWARDS THE SOLUBLE CARBOXYMETHYLCELLULOSE (CMC).

CC -1- FUNCTION: ACTIVE TOWARDS THE SOLUBLE CARBOXYMETHYLCELLULOSE (CMC).

CC -1- FUNCTION: ACTIVE TOWARDS THE SOLUBLE CARBOXYMETHYLCELLULOSE (CMC).

CC -1- FUNCTION: ACTIVE TOWARDS THE SOLUBLE CARBOXYMETHYLCELLULOSE (CMC).

CC -1- FUNCTION: ACTIVE TOWARDS THE SOLUBLE CARBOXYMETHYLCELLULOSE (CMC).

CC -1- FUNCTION: ACTIVE TOWARDS THE SOLUBLE CARBOXYMETHYLCELLULOSE (CMC).

CC -1- FUNCTION: ACTIVE TOWARDS THE SOLUBLE CARBOXYMETHYLCELLULOSE (CMC).

CC -1- FUNCTION: ACTIVE TOWARDS THE SOLUBLE CARBOXYMETHYLCELLULOSE (CMC).

CC -1- FUNCTION: ACTIVE TOWARDS THE SOLUBLE CARBOXYMETHYLCELLULOSE (CMC).

CC -1- FUNCTION: ACTIVE TOWARDS THE SOLUBLE CARBOXYMETHYLCELLULOSE (CMC).

RT "Nucleotide sequences of the envelope genes of two isolates of feline
 RT leukemia virus subgroup B.":
 RL J. Virol. 49:629-632(1984).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=83216322; PubMed=6304347;
 RA Elder J.H., Mullins J.I.;
 RT "Nucleotide sequence of the envelope gene of Gardner-Arnstein feline
 RT leukemia virus B reveals unique sequence homologies with a murine
 RT mink cell focus-forming virus.":
 RL J. Virol. 46:871-880(1983).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=84131936; PubMed=6321156;
 RA Wunsch M., Schulz A.S., Koch W., Friedrich R., Hunsman G.;
 RT "Sequence analysis of Gardner-Arnstein feline leukemia virus
 RT envelope gene reveals common structural properties of mammalian
 RT retroviral envelope genes.":
 RL EMBO J. 2:2239-2246(1983).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88044502; PubMed=2823466;
 RA Guillot S., Hampe A., D'Auriol L., Galibert F.;
 RT "Nucleotide sequence analysis of the LTRs and env genes of SM-FesV
 RT and GA-FesV.":
 RL Virology 161:252-258(1987).
 CC -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: K01209; AAA43052.1; -
 CC EMBL: V01172; CAA24497.1; -
 CC EMBL: X00188; CAA25008.1; -
 CC EMBL: M23026; -; NOT_ANNOTATED_CDS.
 CC PIR: A03991; VCVWGF.
 CC PIR: B33741; VCVW52.
 CC HSSP: P03385; IMOF.
 CC InterPro: IPR002050; Env_polyprotein.
 CC Pfam: PF00429; Env_polyprotein.1.
 CC Coa protein; Glycoprotein; Polyprotein; Signal.
 KW SIGNAL
 FT 1 33
 FT CHAIN 34 465
 FT CHAIN 466 645
 FT CHAIN 646 662
 FT CARBOHYD 43 43
 FT CARBOHYD 58 58
 FT CARBOHYD 286 286
 FT CARBOHYD 322 322
 FT CARBOHYD 327 327
 FT CARBOHYD 351 351
 FT CARBOHYD 354 354
 FT CARBOHYD 394 394
 FT CARBOHYD 410 410
 FT CARBOHYD 430 430
 FT CARBOHYD 430 430
 FT CARBOHYD 15 15
 FT CARBOHYD 41 41
 FT CARBOHYD 47 47
 FT CARBOHYD 51 56
 FT CARBOHYD 70 75
 FT CARBOHYD 80 95
 FT CARBOHYD 99 110
 FT CARBOHYD 120 123
 FT CARBOHYD 127 127
 FT CARBOHYD 134 134
 FT CARBOHYD 143 148
 FT CARBOHYD 158 193

FT FT
 FT CONFLICT 208 208
 FT CONFLICT 215 215
 FT CONFLICT 223 223
 FT CONFLICT 232 232
 FT CONFLICT 238 238
 FT CONFLICT 264 300
 SQ SEQUENCE 662 AA; 73149 MW; 1482088D547CEFF47 CXC64;
 Query Match 7.2%; Score 86.5; DB 1; Length 662;
 Best Local Similarity 20.3%; Pred. No. 2.3;
 Matches 32; Conservative 15; Mismatches 46; Indels 65; Gaps 5;
 QY 14 GKGDSCTMGAGCGYGNLY-----MNGYGLYNALSSALFNDGAMGACGYT 59
 DB 126 GPDGFCAYWGGCTTGTYWRPPSSMDYTYVKKYQGIYQC-----SGGKWCPCPYD 178
 QY 60 IYCDTSQTKWCKRCGNSITITATNLCPMALPNSGCGNCPPLXHFDMSS--QAPAMENIA 117
 DB 179 KAVHSTT-----GASGGRGNLILQFPQKGRQTSMDG-- 212
 QY 118 VYQAGIVPNVKNRVPXORSGRIRFASIGHYELVYVT 155
 DB 213 -----PKSMGLRLXRSGYDPLALFSVS 234
 RESULT 12
 ENV_FSVST STANDARD; PRT; 534 AA.
 ID ENV_FSVST
 AC P03392;
 DT 21-JUL-1986 (rel. 01, Created)
 DT 21-JUL-1986 (rel. 01, Last sequence update)
 DT 16-OCT-2001 (rel. 40, Last annotation update)
 DE ENV_Polyprotein precursor (Coat polyprotein) [contains: knob protein
 DE GP70; Spike protein p15e] (Fragment).
 GN ENV.
 OS feline sarcoma virus (strain Snyder-Theilen).
 OC Viruses; Retroviruses; Retroviridae; Mammalian type C retroviruses.
 OX NCBI_Taxid=11780;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=84115095; PubMed=6319767;
 RA Numburg J.H., Williams M.E., Innis M.A.;
 RT "Nucleotide sequences of the envelope genes of two isolates of feline
 RT leukemia virus subgroup B.":
 RL J. Virol. 49:629-632(1984).
 CC -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: K01208; -; NOT_ANNOTATED_CDS.
 CC PIR: A03992; VCVW5F.
 CC HSSP: P03385; IMOF.
 CC InterPro: IPR002050; Env_polyprotein.
 CC Pfam: PF00429; Env_polyprotein.1.
 CC Coa protein; Glycoprotein; Polyprotein; Signal.
 KW SIGNAL
 FT 1 33
 FT CHAIN 34 465
 FT CHAIN 466 534
 FT CARBOHYD 43 43
 FT CARBOHYD 58 58
 FT CARBOHYD 286 286
 FT CARBOHYD 322 322
 FT CARBOHYD 327 327
 FT CARBOHYD 351 351
 FT CARBOHYD 354 354

Db 128 GYSYHFDI 135

RESULT 9
ENV_FLVC6
ID ENV_FLVC6 STANDARD: PRT: 668 AA.

AC P21443;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE ENV polyprotein precursor [Contains: Coat protein GP70; Coat protein p15E].
OS Feline leukemia provirus (clone CFE-6).
OC Viruses; Retroviral viruses; Retroviridae; Gammaretrovirus.
NCBI_TaxID=11922;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89199802; PubMed=2539525;
RA Kumar D.V., Berry B.T., Roy-Burman P.;
RT "Nucleotide sequence and distinctive characteristics of the env gene of endogenous feline leukemia provirus.";
RL J. Virol. 63:2379-2384(1989).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M25425; AAA30809.1; -
DR PIR: A31479; VCMVFP.
DR HSSP: P03385; IMOF.
DR InterPro: IPR002050; Env.polyprotein.
DR Pfam: PF00429; ENV.polyprotein; 1.
KW Coat protein; Glycoprotein; Polyprotein; Signal.
FT SIGNAL 1 33
FT CHAIN 34 665
FT CHAIN 466 668
FT CARBOHYD 43 43
FT CARBOHYD 58 58
FT CARBOHYD 286 286
FT CARBOHYD 322 322
FT CARBOHYD 327 327
FT CARBOHYD 351 351
FT CARBOHYD 354 354
FT CARBOHYD 394 394
FT CARBOHYD 410 410
FT CARBOHYD 430 430
SQ SEQUENCE 668 AA; 74298 MW; 4A6A06CE2EB8CE25 CRC64;

Query Match
Best Local Similarity 7.6%; Score 91.5; DB 1; Length 668;
Matches 33; Conservative 14; Mismatches 46; Indels 65; Gaps 5;

QY 14 GGRKSGCTMGACGYGNYL-----NAGYGLYNALSSALFNDGAMGACGYT 59
ID ENV_FLVC6 STANDARD: PRT: 668 AA.
AC P03391; P21446;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE ENV polyprotein precursor [Contains: Coat protein GP70; Coat protein p15E; R protein].
OS Feline sarcoma virus (strain Gardner-Arnstein) (Ga-FesV) (Gardner-Arnstein feline leukemia oncovirus B).
OC Viruses; Retroviral viruses; Retroviridae; Gammaretrovirus.
NCBI_TaxID=11774;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84115095; PubMed=6319767;
RA Nunberg J.H., Williams M.E., Innis M.A.;

AC P21444;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE ENV polyprotein precursor [Contains: Coat protein GP70] (Fragment).
OS Feline leukemia provirus (clone CFE-16).
OC Viruses; Retroviral viruses; Retroviridae; Gammaretrovirus.
NCBI_TaxID=11921;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89199802; PubMed=2539525;
RA Kumar D.V., Berry B.T., Roy-Burman P.;
RT "Nucleotide sequence and distinctive characteristics of the env gene of endogenous feline leukemia provirus.";
RL J. Virol. 63:2379-2384(1989).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M25582; AAA30810.1; -
DR PIR: B31479; B31479.
DR HSSP: P03390; IACF.
DR InterPro: IPR002050; Env.polyprotein.
DR Pfam: PF00429; ENV.polyprotein; 1.
KW Coat protein; Glycoprotein; Polyprotein; Signal.
FT SIGNAL 1 33
FT CHAIN 34 >273
FT CARBOHYD 43 43
FT CARBOHYD 58 58
FT CARBOHYD 273 273
SQ SEQUENCE 273 AA; 30008 MW; 6B605FCD582B325D CRC64;

Query Match
Best Local Similarity 7.2%; Score 86.5; DB 1; Length 273;
Matches 32; Conservative 15; Mismatches 46; Indels 65; Gaps 5;

QY 14 GGRKSGCTMGACGYGNYL-----NAGYGLYNALSSALFNDGAMGACGYT 59
ID ENV_FSVGA STANDARD: PRT: 662 AA.
AC P03391; P21446;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE ENV polyprotein precursor [Contains: Coat protein GP70; Coat protein p15E; R protein].
OS Feline sarcoma virus (strain Gardner-Arnstein) (Ga-FesV) (Gardner-Arnstein feline leukemia oncovirus B).
OC Viruses; Retroviral viruses; Retroviridae; Gammaretrovirus.
NCBI_TaxID=11774;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84115095; PubMed=6319767;
RA Nunberg J.H., Williams M.E., Innis M.A.;


```

ID MP21_MAIZE STANDARD; PRT; 191 AA.
AC 007154;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DE 01-MAR-2002 (Rel. 41, Last annotation update)
DE Pollen allergen Zea m 1 (Zea m 1).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pollen;
RX MEDLINE=94010312; PubMed=8406014;
RA Broadwater A.H., Rubinstein A.L., Chay C.H., Klapper D.G.,
RA Bedinger P.A.;
RT "Zea m1, the maize homolog of the allergen-encoding Lol p1 gene of
RT ryegrass."
RL Gene 131:227-230(1993).
CC -1- TISSUE SPECIFICITY: POLLEN TISSUE.
CC -1- DEVELOPMENTAL STAGE: EXPRESSION LOW BEFORE AND HIGH AFTER
CC POLLEN MITOSIS.
CC -1- DISEASE: CAUSES MAIZE POLLEN ALLERGY.
CC -1- SIMILARITY: BELONGS TO THE LOL P 1 FAMILY OF ALLERGENS.
CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE EG45 DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE CBD DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: L14271; AAA33496.1; -
DR PIR: JC1524; JC1524.
DR HSSP: P43214; IMHO.
DR MalzEDB: 65840; -
DR InterPro: IPR000882; Pollen_allergen.
DR Pfam: PF01357; Pollen_allergen.1.
DR PRINTS: PR01225; EXPANSIN_FAMLY.
DR ProDom: PD002179; Pollen_allergen.1.
DR PROSITE: PS50843; EXPANSIN_CBD; 1.
DR PROSITE: PS50842; EXPANSIN_EG45; 1.
KW Allergen; multigene family.
FT DOMAIN 1 91 EXPANSIN-LIKE EG45.
FT DOMAIN 105 186 EXPANSIN-LIKE CBD.
SQ SEQUENCE 191 AA; 21362 MW; 6E2A9DF921C45C63 CRC64;

Query Match 13.2%; Score 157.5; DB 1; Length 191;
Best Local Similarity 27.7%; Pred. No. 3.4e-07;
Matches 53; Conservative 32; Mismatches 67; Indels 39; Gaps 11;

```

```

ID MPCL1_CYNDA STANDARD; PRT; 246 AA.
AC 004701;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 01-MAR-2002 (Rel. 41, Last annotation update)
DE Major pollen allergen Cyn d 1.
CN CYN d1.
OS Cyndon dactylon (Bermuda grass).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Chloridoideae; Cyndonatae; Cyndon.
OX NCBI_TaxID=28909;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Pollen;
RX MEDLINE=96347957; PubMed=8757211;
RA Smith P.M., Suphloglu C., Griffith I.J., Thierault K., Knox R.B.,
RA Singh M.B.;
RT "Cloning and expression in yeast Pichia pastoris of a biologically
RT active form of Cyn d 1, the major allergen of Bermuda grass pollen."
RL J. Allergy Clin. Immunol. 98:331-343(1996).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DISEASE: CAUSES GRASS POLLEN ALLERGY.
CC -1- SIMILARITY: BELONGS TO THE LOL P 1 FAMILY OF ALLERGENS.
CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE EG45 DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE CBD DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: S83343; AAB50734.2; -
DR HSSP: P43214; IMHO.
DR InterPro: IPR000882; Pollen_allergen.
DR Pfam: PF01357; Pollen_allergen.1.
DR PRINTS: PR01225; EXPANSIN_FAMLY.
DR ProDom: PD002179; Pollen_allergen.1.
DR PROSITE: PS50843; EXPANSIN_CBD; 1.
DR PROSITE: PS50842; EXPANSIN_EG45; 1.
KW Allergen.
FT DOMAIN 39 145 EXPANSIN-LIKE EG45.
FT DOMAIN 159 240 EXPANSIN-LIKE CBD.
FT CARBOHYD 9 N-LINKED (GLCNAC... ) (POTENTIAL).
SQ SEQUENCE 246 AA; 26888 MW; 43DB8442DBA588322 CRC64;

Query Match 13.1%; Score 157; DB 1; Length 246;
Best Local Similarity 25.5%; Pred. No. 4.8e-07;
Matches 59; Conservative 38; Mismatches 96; Indels 38; Gaps 11;

```

RESULT 6

RESULT 7

RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-ANther;
 RX MEDLINE=96069591; PubMed=7590339;
 RA Xu H., Theerakulpisut P., Goulding N., Suphloglu C., Singh M.B.,
 RT Bhalla P.L.,
 RT Cloning, expression and immunological characterization of Ory s 1,
 RT the major allergen of rice pollen.
 RL Gene 164:255-259(1995).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN MATURE ANTHERS BUT NOT IN
 CC VEGETATIVE OR OTHER FLORAL TISSUES.
 CC -1- DISEASE: CAUSES GRASS POLLEN ALLERGY.
 CC -1- SIMILARITY: BELONGS TO THE LOL P I FAMILY OF ALLERGENS.
 CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE EG45 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE CBD DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 CC EMBL: U31771; AAA86533.1; -
 CC HSSP: P43214; IMHO.
 CC InterPro: IPR000882; Pollen_allergen.
 CC Pfam: PF01357; Pollen_allergen; 1.
 CC PRINTS: PR01225; EXPANSINFAMILY.
 CC ProDom: PD002179; Pollen_allergen; 1.
 CC PROSITE: PS50843; EXPANSIN_CBD; 1.
 CC PROSITE: PS50842; EXPANSIN_EG45; 1.
 CC Allergen: Glycoprotein; Signal; Multigene family.
 CC FT SIGNAL 1 23 POTENTIAL.
 CC FT CHAIN 24 263 MAJOR POLLEN ALLERGEN ORY S 1.
 CC FT DOMAIN 61 164 EXPANSIN-LIKE EG45.
 CC FT DOMAIN 178 259 EXPANSIN-LIKE CBD.
 CC FT CARBOHYD 32 32 N-LINKED (GLCNAC...) (POTENTIAL).
 CC SQ SEQUENCE 263 AA; 28497 MW; B1C5F24EA398DD60 CRC64;

 CC Query Match 13.7%; Score 163.5; DB 1; Length 263;
 CC Best Local Similarity 29.8%; Pred. No. 1.4e-07;
 CC Matches 53; Conservative 29; Mismatches 63; Indels 33; Gaps 9;

 QY 23 GGACGYNLYNAGYGLYNAALSSALENDGAMGACATYTTCDTSQTKWCKRGSGNSTTTAT 82
 DB 61 GGACGYKDVDPKAPLGMNSGNDPIFKDGRGCGSCFEIKC--SKPEACSDKPALIHVTDM 118
 QY 83 NLCKPNNALPSNSGWCNPLX--HFDMSQPAW---ENIAYVOAGVAPVNYKRPVXORS 136
 DB 119 N-----DEPLAAYHFDLSGLAAMAKDKDELRKAGIITDQFRVRYCKYP 162
 QY 137 GG--IRFAI---SGHDYFELVYTNNGSGVVAQMSI--KSGNTGMAMSRNMGANMO 187
 DB 163 ADTITTHIEKASNPVYLL--VKYVAGDGDVVEVEIKEKGSSE--WALKESMGAIWR 218

 RESULT 4
 MPRL_HOLLA STANDARD; PRT; 265 AA.
 AC P43216; 039975;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Major pollen allergen Hol 1 precursor (Hol 1 I) (Hol 1 I.0101 and
 DE 1.0102).
 OS Holcus lanatus (Velvet grass).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
 OC Poaceae; Holcus.
 OC NCBI_TaxID=29679;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN-CV. AVENEAE; TISSUE=Pollen;
 RA Schramm G.D., Bufe A., Becker W.M., Schlaak M.;
 RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 18-265 FROM N.A.
 RC STRAIN-CV. AVENEAE; TISSUE=Pollen;
 RX MEDLINE=97358126; PubMed=9215246;
 RA Schramm G.D., Bufe A., Petersen A., Haas H., Schlaak M., Becker W.M.;
 RT Mapping of IgE-binding epitopes on the recombinant major group I
 RT allergen of velvet grass pollen, Hol 1 I.
 RL J. Allergy Clin. Immunol. 99:781-787(1997).
 RN [3]
 RP CHARACTERIZATION
 RC STRAIN-CV. AVENEAE; TISSUE=Pollen;
 RX MEDLINE=96319506; PubMed=8768803;
 RA Schramm G.D., Petersen A., Bufe A., Schlaak M., Becker W.M.;
 RT Identification and characterization of the major allergens of velvet
 RT grass (Holcus lanatus), Hol 1 I and Hol 1 I.
 RL Int. Arch. Allergy Immunol. 110:354-363(1996).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO THE LOL P I FAMILY OF ALLERGENS.
 CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE EG45 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE CBD DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 CC EMBL: Z27084; CA81610.1; -
 CC EMBL: Z68893; CA93121.1; -
 CC HSSP: P43214; IMHO.
 CC InterPro: IPR000882; Pollen_allergen.
 CC Pfam: PF01357; Pollen_allergen; 1.
 CC PRINTS: PR01225; EXPANSINFAMILY.
 CC ProDom: PD002179; Pollen_allergen; 1.
 CC PROSITE: PS50843; EXPANSIN_CBD; 1.
 CC PROSITE: PS50842; EXPANSIN_EG45; 1.
 CC Allergen: Signal.
 CC FT SIGNAL 1 25 POTENTIAL.
 CC FT CHAIN 26 265 MAJOR POLLEN ALLERGEN HOL 1 I.
 CC FT DOMAIN 63 169 EXPANSIN-LIKE EG45.
 CC FT DOMAIN 183 264 EXPANSIN-LIKE CBD.
 CC FT VARIANT 103 103 T->S (IN HOL 1 I.0102).
 CC SQ SEQUENCE 265 AA; 28590 MW; 1FE23B3EE198AD6D CRC64;

 CC Query Match 13.3%; Score 159.5; DB 1; Length 265;
 CC Best Local Similarity 26.2%; Pred. No. 3.1e-07;
 CC Matches 61; Conservative 39; Mismatches 90; Indels 43; Gaps 13;

 QY 8 AFATFYGKRGDSCGM--GGACGYNLYNAGYGLYNAALSSALEPNDGAMGACATYTTCDMS 65
 DB 46 AKSTWYKPTGAGPKDNGGACGYKDVDPFSGMTGCGTPIFKGRGCGSCFEIKC--- 102
 QY 66 QTKWCKRGSGNSTITTTATNLCXPMWALPSNSGWCNPLX--HFDMSQPAWMI----- 117
 DB 103 -TRPESCGEPVYVHTD-----DN-----EPLIAPYHFDLSGHNFGSAKKGEQ 147
 QY 118 -VVOAGIVPYNRYKRPVXORSGR--FAI---SGHDYFELVYTNNGSGVVAQMSIKGS 171
 DB 148 KLSAGLELEKFRVCKRKYDGTGKPFVHEKGSNPVYLL--VKYLDGGDVVAVDIKRK 206
 QY 172 NTG--WMAMSRNMGANQMSNA--YLAQGSLSFTVO--LDDGRKVTANNAKPPXW 219
 DB 207 GKDKWIELEKSGWAVWRVDPDKLTG---PFTVRYTTEGTGAEADVPEGM 256

 RESULT 5
 MP21_MAIZE

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: M57474; AAA63279.1; -
 DR EMBL: M57476; AAA63278.1; -
 DR PIR: A23341; A23341.
 DR PIR: B37881; B37881.
 DR PIR: S13614; S13614.
 DR HSP: P43214; IMHO.
 DR InterPro: IPR000882; Pollen_allergen.
 DR Pfam: PF01357; Pollen_allergen. 1.
 DR PRINTS: PR01225; EXPANSIN-FAMILY.
 DR Prodom: PD002179; Pollen_allergen. 1.
 DR PROSITE: PS50843; EXPANSIN-CBD; 1.
 DR PROSITE: PS50842; EXPANSIN-EG45; 1.
 DR Allergen: Glycoprotein; Signal; Multigene family.
 FT SIGNAL 1 23
 FT CHAIN 24 263 POLLEN ALLERGEN LOI P 1.
 FT DOMAIN 61 167 EXPANSIN-LIKE EG45.
 FT CARBOHYD 181 262 EXPANSIN-LIKE CBD.
 FT VARIANT 32 32 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARIANT 68 68 N -> D.
 FT VARIANT 177 177 A -> G.
 FT VARIANT 210 210 I -> T.
 FT VARIANT 246 246 F -> V.
 FT CONFLICT 28 28 P -> C (IN REF. 3).
 FT CONFLICT 31 31 P -> C (IN REF. 3).
 FT CONFLICT 48 48 W -> MW (IN REF. 3).
 SQ SEQUENCE 263 AA; 28438 MW; 7675896F29C88C9 CRC64;

Query Match 14.4%; Score 172.5; DB 1; Length 263;
 Best Local Similarity 26.3%; Pred. No. 2.2e-08;
 Matches 61; Conservative 36; Mismatches 94; Indels 41; Gaps 11;

QY 8 AFATFYGGKDGSCM--GGACGYGLNAGYGLYNAALSFNDGAMCGACTYTTCDTS 65
 D 4 AKSTWYKPTGAPKPDNGACGYKNDKAPFNGMTGGNTPIFDGCGSCGFELKC--- 100
 QY 66 QTKWCKPGNSITTTATNLCXPNMALPSNSGCMCNPLX--HDMSPAMENIA----- 117
 D 101 -TKPSCSGEAVTVTTD-----DN-----EPIAPYHFDLSGHAFSGMAKKEEQ 145
 QY 118 -VYAGIVPVNKKYKRR--PXQSGGIRFALSGHDFELVTVTVGSGVYAAQMSIGSN 172
 D 146 NVKAGLELQFRVKKCKYPPDPTPTFHEKASNPNTLALVTVDDGDVAVADIEKG 205
 QY 173 TG-WMAMSRNMGAMQNSA--YLAGQSLSTVQ--LDDGRKVTAMNXPXNM 219
 D 206 KDKWIELKESGAVWIRIDTPDKLTG---PFTVRYTTEGTRKSEFEDYIPGCM 254

RESULT 2
 MPAL_PHAQO STANDARD; PRT; 269 AA.
 ID MPAL_PHAQO
 AC 041260;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Major pollen allergen Pha a 1 precursor (Pha a 1).
 OS Phalaris aquatica (Canary grass).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliopsida; Liliopsida; Poales; Poaceae; Poideae;
 OC Poaceae; Phalaris.
 NCBI_TaxID=28479;
 RX NCBI_TaxID=28479;
 RP SEQUENCE FROM N.A.
 RC TISSUE=pollen;
 RX MEDLINE=96105369; PubMed=8564724;
 RA Suphioglu C., Singh M.B.;
 RT "Cloning, sequencing and expression in Escherichia coli of Pha a 1
 and four isoforms of Pha a 5, the major allergens of canary grass
 pollen.";

RL Clin. Exp. Allergy 25:853-865(1995).
 RN [2]
 RP SEQUENCE OF 30-49.
 RX MEDLINE=93319091; PubMed=7687099;
 RA Suphioglu C., Singh M.B., Simpson R.J., Ward L., Knox R.B.;
 RT "Identification of canary grass (Phalaris aquatica) pollen allergens
 by immunoblotting: IgE and IgG antibody-binding studies.";
 RL Allergy 48:273-281(1993).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- DISEASE: CAUSES GRASS POLLEN ALLERGY.
 CC -1- SIMILARITY: BELONGS TO THE LOI P 1 FAMILY OF ALLERGENS.
 CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE EG45 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE CBD DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

DR EMBL: S80654; AAB35984.1; -
 DR HSP: P43214; IMHO.
 DR InterPro: IPR000882; Pollen_allergen.
 DR Pfam: PF01357; Pollen_allergen. 1.
 DR PRINTS: PR01225; EXPANSIN-FAMILY.
 DR Prodom: PD002179; Pollen_allergen. 1.
 DR PROSITE: PS50843; EXPANSIN-CBD; 1.
 DR PROSITE: PS50842; EXPANSIN-EG45; 1.
 DR Allergen: Glycoprotein; Signal.
 FT SIGNAL 1 29
 FT CHAIN 30 269 MAJOR POLLEN ALLERGEN PHA A 1.
 FT DOMAIN 67 173 EXPANSIN-LIKE EG45.
 FT DOMAIN 187 268 EXPANSIN-LIKE CBD.
 FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 37 37 P -> G (IN REF. 2).
 SQ SEQUENCE 269 AA; 29011 MW; 058834A58ACEB77F CRC64;

Query Match 13.7%; Score 164.5; DB 1; Length 269;
 Best Local Similarity 27.0%; Pred. No. 1.1e-07;
 Matches 63; Conservative 39; Mismatches 88; Indels 43; Gaps 14;

QY 8 AFATFYGGKDGSCM--GGACGYGLNAGYGLYNAALSFNDGAMCGACTYTTCDTS 65
 D 50 AKSTWYKPTGAPKPDNGACGYKNDKAPFNGMTGGNTPIFDGCGSCGFELKC--S 107
 QY 66 QTKWCKPGNSITTTATNLCXPNMALPSNSGCMCNPLX--HDMSPAMENIA----- 117
 D 108 KPESC--SGEPTVHTD-----DN-----EPIAPYHFDLSGHAFSGMAKKEEQ 151
 QY 118 -VYAGIVPVNKKYKRR--PXQSGGIR--FAL--SGHDFELVTVTVGSGVYAAQMSIGSN 171
 D 152 NVKAGLELQFRVKKCKYPPDPTPTFHEKASNPNTLALVTVDDGDVAVADIEK 210
 QY 172 NMG-WMAMSRNMGAMQNSA--YLAGQSLSTVQ--LDDGRKVTAMNXPXNM 219
 D 211 GDKWIELKESGAVWIRIDTPDKLTG---PFTVRYTTEGTRKSEFEDYIPGCM 260

RESULT 3
 MPOL_ORYSA STANDARD; PRT; 263 AA.
 ID MPOL_ORYSA
 AC 040638;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Major pollen allergen Ory s 1 precursor (Ory s 1).
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliopsida; Liliopsida; Poales; Poaceae;
 OC Euphorbiaceae; Oryzaeae; Oryza.
 NCBI_TaxID=4530;

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 11, 2002, 14:50:30 ; Search time 5.86347 Seconds
(without alignments)
1499.000 Million cell updates/sec

Title: US-09-896-301-4
Perfect score: 1197
Sequence: 1 KXSVAGSFAFYGKDGSC.....KVTAMNXPXNMLXXXXXX 227

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	172.5	14.4	269	1 MP1L_LOLPR	P14946 lolium pere
2	164.5	13.7	269	1 MP1L_PHAPO	P41260 phalaris ag
3	163.5	13.7	263	1 MP1L_ORYSA	O40638 oryza sativ
4	159.5	13.3	265	1 MP1L_HOLIA	P43216 holcus lana
5	157.5	13.2	191	1 MP1L_MAIZE	Q07154 zea mays (m
6	157.5	13.1	246	1 MP1L_CYNDA	O04701 cynodon dac
7	152.5	12.7	263	1 MP1L_PHLPR	P43213 phleum prat
8	134	11.2	242	1 GUN5_TIRRE	P43317 trichoderma
9	91.5	7.6	668	1 ENY_FLV6	P21443 feline leuk
10	86.5	7.2	273	1 ENY_FLVCI	P03391 feline sarc
11	86.5	7.2	662	1 ENY_FSVGA	P03392 feline sarc
12	85.5	7.1	534	1 ENY_FSVST	P11261 feline leuk
13	85.5	7.1	662	1 ENY_FLVLB	P82186 mytilus edu
14	85	7.1	181	1 GUN_MYTE	P11370 mus musculi
15	84.5	7.1	432	1 ENY2_MOUSE	P43316 humicola in
16	82.5	6.9	213	1 GUN5_HUMIN	O9X1E3 thermotoga
17	82	6.9	234	1 GLPF_THEMA	Q02077 feline leuk
18	82	6.9	436	1 ENY_FLVCS	P03455 influenza a
19	81.5	6.8	566	1 HEMA_IJZNT	O9Y5X1 homo sapien
20	81	6.7	595	1 SNX9_HUMAN	P80468 struthio ca
21	80.5	6.7	379	1 ADH2_SMRCA	O59149 anabaena sp
22	80	6.7	662	1 PRCA_ANASP	P82474 zingiber of
23	79	6.6	221	1 GP1T_ZINOF	P53067 ratius norv
24	79	6.6	264	1 GLPF_STRCO	P19255 streptomyce
25	78.5	6.6	329	1 CYST_CAEEL	P25807 caenorhabdi
26	78.5	6.6	1385	1 YMSI_CAEEL	P34501 caenorhabdi
27	78.5	6.6	1268	1 PCGN_MOUSE	P55066 mus musculi
28	77.5	6.5	433	1 SH06_ECOLI	P09750 escherichia
29	77.5	6.5	444	1 SH05_ECOLI	P09749 escherichia
30	77.5	6.5	444	1 SCAB_XENLA	P51169 xenopus lae
31	77.5	6.5	667	1 TSPE_BPP22	P12528 bacterioph
32	77.5	6.5	714	1 DL1L_RAT	P97677 ratius norv
33	77.5	6.5	714	1 DL1L_RAT	P97677 ratius norv

34	77	6.4	348	1 PAP4_CARPA	P05994 carica papa
35	77	6.4	362	1 PCGY_CHICK	O90953 gallus gall
36	76.5	6.4	360	1 CYSE_HBMS	P43156 hemerocalli
37	76	6.3	405	1 VGLG_HSV4	P32650 equine herp
38	76	6.3	410	1 ENY_FLVCA	Q02076 feline leuk
39	76	6.3	456	1 SHU7_ECOLI	P09751 escherichia
40	76	6.3	493	1 YBGH_ECOLI	P75742 escherichia
41	76	6.3	642	1 ENY_FLVGL	P08359 feline leuk
42	76	6.3	1045	1 GUNB_CHLFI	P26325 cellulomona
43	76	6.3	1822	1 ITB4_HUMAN	P16144 homo sapien
44	75.5	6.3	82	1 KRHA_SHEEP	Q02958 ovis aries
45	75.5	6.3	573	1 ALP1_YEAST	P38971 saccharomyc

ALIGNMENTS

RESULT 1

MP1L_LOLPR STANDARD; PRT; 263 AA.
ID MP1L_LOLPR
AC P14946; P19964;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Pollen allergen lol p 1 precursor (lol p I) (Allergen R7).
OS Lolium perenne (Perennial ryegrass).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Poaceae; Lolium.
NCBI_TaxID=4522;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90375479; PubMed=1697854;
RA Perez M., Ishioka G.Y., Walker L.E., Chesnut R.W.;
RT "cDNA cloning and immunological characterization of the rye grass
allergen lol p I.";
RL J. Biol. Chem. 265:16210-16215(1990).
RN [2]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 24-48.
RC TISSUE=Pollen;
RX MEDLINE=91160716; PubMed=2001733;
RA Griffith I.J., Smith P.M., Pollock J., Theerakulpisut P.,
RA Avoglu A., Davies S., Hough T., Singh M.B., Simpson R.J., Ward I.D.,
RA Knox R.B.;
RT "Cloning and sequencing of lol pI, the major allergenic protein of
rye grass pollen.";
RL FEBS Lett. 279:210-215(1991).
RN [3]
RP SEQUENCE OF 24-53.
RC TISSUE=Pollen;
RX MEDLINE=86242068; PubMed=3718469;
RA Cottam G.P., Moran D.M., Strandberg R.;
RT "Physicochemical and immunochemical characterization of allergenic
proteins from rye-grass (Lolium perenne) pollen prepared by a rapid
and efficient purification method.";
RL Biochem. J. 234:305-310(1986).
RN [4]
RP SEQUENCE OF 236-263.
RX MEDLINE=89364850; PubMed=2475768;
RA Esch R.E., Klapper D.G.;
RT "Isolation and characterization of a major cross-reactive grass group
I allergenic determinant.";
RL Mol. Immunol. 26:557-561(1989).
RN [5]
RP SUBCELLULAR LOCATION: Secreted.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DISEASE: CAUSES GRASS POLLEN ALLERGY.
CC -1- SIMILARITY: BELONGS TO THE LOL P I FAMILY OF ALLERGENS.
CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE EC45 DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE CBD DOMAIN.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way

Query Match	63.58;	Score 759.5;	DB 10;	Length 253;
Best Local Similarity	63.38;	Pred. NO. 3.5e-62;		
Matches`_333;	Conservative 31;	Mismatches 45;	Indels 1;	Gaps 1

Query Match	63.48;	Score 759;	DB 10;	Length 252;
Best Local Similarity	65.08;	Pred. No. 3.9e-62;		
Matches 139;	Conservative 24;	Mismatches 49;	Indels 2;	Gaps 1;

```
QY      186 WQSNAYLAGQSLFTVQLDDGRKVTANMXAPXNW 219
        ||||| | ||||| | : ||| : : | | | |
Db      208 WQSNADLDGQSLSEKVTISDGRIVSNNAAPACW 241
```

Search completed: October 11, 2002, 14:59:41
Job time : 20.6007 secs

RL Submitted (SEP-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AC007789; AAD38297.1; -
 DR EMBL: AP002865; BAB18338.1; -
 DR InterPro: IPR000882; Pollen_allergen.
 DR Pfam: PF01357; Pollen_allergen; 1.
 DR PRINTS: PR01225; EXPANSNFAMILY.
 DR ProDom: PD002179; Pollen_allergen; 1.
 DR SEQUENCE 254 AA; 26987 MW; A40924E21A819101 CRC64;

Query Match 67.6%; Score 809.5; DB 10; Length 254;
 Best Local Similarity 67.6%; Pred. No. 8.9e-67;
 Matches 144; Conservative 22; Mismatches 46; Indels 1; Gaps 1;

OY 7 SAFATFYGGKDSCTMGACGYNLYNAGYGLYNALSSALFENDGAMGACACTTTCDSQ 66
 DB 29 SATATFYGGSDASGTMGSCGCGNMTSAGYGTNTTALSALYGDASGACACTVTCDA 88
 OY 67 TWCKPFGNSITTTATNLCPNMLPNSGWCNPELXHFDM SOPAMENIAYOAGIYPV 126
 DB 89 TWCK-NGTSYVVTATNYCPNYSBGDAGWCNPRHFDMSOPAMENIAYSSGIYPV 147
 OY 127 NKRRPXRSGIRRAISGHDFELYVTYNNVGGSGVVAOMSTKSGNTGMAMSRNMGANW 186
 DB 148 RIAKTPCRRVGGIRFGIAGHDYELVLTYNVAGSAAVAAMVKGSGTEWLSMRNMGANW 207
 OY 187 QSNATLAGQSLSFIYQDDGRKVTAMNXPXNW 219
 DB 208 QSNATLTGQALSFYQADGCVVTAIVDAVPANW 240

RESULT 5
 Q39625 PRELIMINARY; PRT; 250 AA.

ID Q39625; AC Q39625; DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE EXPANSIN SL.
 GN CS-EXPL.
 OS Cucumis sativus (Cucumber).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Cucurbitales; Cucurbitaceae; Cucumis.
 OX NCBI_TaxID=3659;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BURPEE PICKLER;
 RX MEDLINE=96016146; PubMed=7568110;
 RA Shcherban T.Y., Shi Y., Durachko D.M., Gullinan M.J.,
 RA McQueen-Mason S.J., Shieh M., Cosgrove D.J.;
 RT "Molecular cloning and sequence analysis of expansins--a highly
 RT conserved, multigene family of proteins that mediate cell wall
 RT extension in plants.";
 RL Proc. Natl. Acad. Sci. U.S.A. 92:9245-9249(1995).
 DR EMBL: U30382; AAB37746.1; -
 DR InterPro: IPR000882; Pollen_allergen.
 DR Pfam: PF01357; Pollen_allergen; 1.
 DR PRINTS: PR01225; EXPANSNFAMILY.
 DR ProDom: PD002179; Pollen_allergen; 1.
 FT CHAIN 24 250
 SQ SEQUENCE 250 AA; 27215 MW; 60651BC47E4186DA CRC64;

Query Match 64.6%; Score 773.5; DB 10; Length 250;
 Best Local Similarity 64.5%; Pred. No. 1.8e-63;
 Matches 138; Conservative 27; Mismatches 46; Indels 3; Gaps 2;

OY 6 QSAFATFYGGKDSCTMGACGYNLYNAGYGLYNALSSALFENDGAMGACACTTTCDS 65
 DB 29 QSGHATFYGGSDASGTMGACGYNLYSOGYGTNTTALSALFENNGISGACACEMTC-TN 87
 OY 66 OTKWCPRGNSITTTATNLCPNMLPNSGWCNPELXHFDM SOPAMENIAYOAGIYPV 125
 DB 88 DPAWCLPG-TIRVATINCPNFPALPNNNGWCNPPLOHFDMAEPALFOIAQYRAGIYV 145

OY 126 VVKRVPXRSGIRRAISGHDFELYVTYNNVGGSGVVAOMSTKSGNTGMAMSRNMGAN 185
 DB 146 VSFRVPCMKKGVAFITNGHSYFNLVLTINVGAGDVHSVSTIGSRIGWMSMRNMGAN 205
 OY 186 WOSNATYLAGQSLSFIYQDDGRKVTAMNXPXNW 219
 DB 206 WOSNATYLAGQSLSFIYQVTLSDGRKTLTAVNLVPSNW 239

RESULT 6
 Q94619 PRELIMINARY; PRT; 250 AA.

ID Q94619; AC Q94619; DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE ALPHA-EXPANSIN.
 GN EXP12.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Eupharoidae; Oryzae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21521099; PubMed=11641069;
 RA Lee Y., Choi D., Kende H.;
 RT "Expansins: ever-expanding numbers and functions.";
 RL Curr. Opin. Plant Biol. 4:527-532(2001).
 DR EMBL: AF394548; AAL24484.1; -
 SQ SEQUENCE 250 AA; 26429 MW; AEE0729BAE81079B CRC64;

Query Match 63.1%; Score 772.5; DB 10; Length 250;
 Best Local Similarity 63.1%; Pred. No. 2.2e-63;
 Matches 137; Conservative 25; Mismatches 54; Indels 1; Gaps 1;

OY 3 SVAQSAFATFYGGKDSCTMGACGYNLYNAGYGLYNALSSALFENDGAMGACACTTTC 62
 DB 22 AALSGTATFYGGSDASGTMGACGYNLYSTGYGTNTTALSALFENDGACGCEYQITC 81
 OY 63 DTGOTKWCPRGNSITTTATNLCPNMLPNSGWCNPELXHFDM SOPAMENIAYOAG 122
 DB 82 DQNSKWKCK-AGTSVITTTATNLCPNMLPNSGWCNPELXHFDM SOPAMENIAYOAG 140
 OY 123 IVDVNTKRVXPXRSGIRRAISGHDFELYVTYNNVGGSGVVAOMSTKSGNTGMAMSRN 182
 DB 141 IVDVNTKRVXPXRSGIRRAISGHDFELYVTYNNVGGSGVVAOMSTKSGNTGMAMSRN 200
 OY 183 GAWQSNATYLAGQSLSFIYQDDGRKVTAMNXPXNW 219
 DB 201 GAWQSNATYLAGQSLSFIYQVTLSDGRKTLTAVNLVPSNW 237

RESULT 7
 Q94611

ID Q94611; AC Q94611; PRELIMINARY; PRT; 251 AA.

DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE ALPHA-EXPANSIN.
 GN EXP2.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Eupharoidae; Oryzae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21521099; PubMed=11641069;
 RA Lee Y., Choi D., Kende H.;
 RT "Expansins: ever-expanding numbers and functions.";

QY 61 TCDTSQTKKCKPGNSGCGTTCATATNLCXPNMALPSNSGCGCNPPLXHPDMSOPAMENIAVYQ 120
 DB 84 TCDTSQTKKCKPGNSGCGTTCATATNLCXPNMALPSNSGCGCNPPLXHPDMSOPAMENIAVYQ 143
 QY 121 AGIYPVYKRVKXQKSGIRFPAISGHDYFELVYTVTNVGGSGVVAOMSIKSNMGWAMSR 180
 DB 144 AGIYPVYKRVKXQKSGIRFPAISGHDYFELVYTVTNVGGSGVVAOMSIKSNMGWAMSR 203
 QY 181 MCGAMQSNAYLAGOSLSFTVQDDGKRVYAMNXPXNM 219
 DB 204 MCGAMQSNAYLAGOSLSFTVQDDGKRVYAMNXPXNM 242

RESULT 2
 Q9XHX0 PRELIMINARY: PRT: 251 AA.
 ID Q9XHX0
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE PUTATIVE EXPANSIN.
 GN T49B20.23 OR P0034C11.27.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Euphorbiaceae; Oryzaeae; Oryza.
 NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. NIPPONBARE;
 RA Buell C.R., Benito M.-I., Lin X., Mason T.M., Umayam L., Shea T.P.,
 RA Fujii C.Y., Shen M., Fraser C.M.;
 RT "Oryza sativa BAC T49B20 genomic sequence."
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
 clone:P0034C11."
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC007789; AAD38296.1; -
 DR EMBL: AP002865; BAB1836.1; -
 DR InterPro: IPR000882; Pollen_allergen.1.
 DR Pfam: PF01357; Pollen_allergen.1.
 DR PRINTS: PR01225; EXPANSINFAMLY.
 DR ProDom: PD002179; Pollen_allergen.1.
 DR ProDom: PD002179; Pollen_allergen.1.
 SQ SEQUENCE 251 AA; 26334 MW; 1A0BB36709FDF9A3 CRC64;

Query Match 74.0%; Score 885.5; DB 10; Length 251;
 Best Local Similarity 76.4%; Pred. No. 8.8e-74;
 Matches 162; Conservative 15; Mismatches 34; Indels 1; Gaps 1;
 QY 8 AFATFYGGKDGSGCTMGACGCGNLYNAGLYNAAALSSALFNDGAMCGACYTTCDSOT 67
 DB 28 AFATFYGGKDGSGCTMGACGCGNLYNAGLYNAAALSSALFNDGAMCGACYTTCDSOT 87
 QY 68 KWCKPGNSGIRFPAISGHDYFELVYTVTNVGGSGVVAOMSIKSNMGWAMSR 127
 DB 88 QMCKP-GTSTVTATNCECPNVALSGDAGWCNPRRHEDMSQPAWETIAVYRAGIYV 146
 QY 128 YKRVKXQKSGIRFPAISGHDYFELVYTVTNVGGSGVVAOMSIKSNMGWAMSR 187
 DB 147 YKRVKXQKSGIRFPAISGHDYFELVYTVTNVGGSGVVAOMSIKSNMGWAMSR 206
 QY 188 SNAVLAGOSLSFTVQDDGKRVYAMNXPXNM 219
 DB 207 SNAVLAGOSLSFTVQDDGKRVYAMNXPXNM 238

RESULT 3
 Q9FY29 PRELIMINARY: PRT: 255 AA.
 ID Q9FY29

AC Q9FY29;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE EXPANSIN.
 GN EXP3.
 OS Festuca pratensis.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
 OC Poaeae; Festuca.
 NCBI_TaxID=4608;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. PREEST; TISSUE=ROOT;
 RX MEDLINE=2137433; PubMed=11485205;
 RA Reidy B., McQueen-Mason S., Noesberger J., Fleming A.;
 RT "Differential Expression of alpha- and beta-expansin genes in the
 RL elongating leaf of Festuca pratensis."
 RL Plant Mol. Biol. 46:491-504(2001).
 DR EMBL: AJ276008; CAC06434.1; -
 DR InterPro: IPR000882; Pollen_allergen.
 DR PRINTS: PR01225; EXPANSINFAMLY.
 DR ProDom: PD002179; Pollen_allergen.1.
 SQ SEQUENCE 255 AA; 27358 MW; 5B1082946042B3DC CRC64;

Query Match 72.4%; Score 866.5; DB 10; Length 255;
 Best Local Similarity 74.6%; Pred. No. 5e-72;
 Matches 159; Conservative 18; Mismatches 35; Indels 1; Gaps 1;
 QY 7 SAFATFYGGKDGSGCTMGACGCGNLYNAGLYNAAALSSALFNDGAMCGACYTTCDSOT 66
 DB 31 SAFATFYGGKDGSGCTMGACGCGNLYNAGLYNAAALSSALFNDGAMCGACYTTCDSOT 90
 QY 67 TKWCKPGNSGIRFPAISGHDYFELVYTVTNVGGSGVVAOMSIKSNMGWAMSR 126
 DB 91 SMCN-AGKSTITSATNLCXPNMALPSNSGCGCNPPLXHPDMSOPAMENIAVYQAGIYV 149
 QY 127 NKRVKXQKSGIRFPAISGHDYFELVYTVTNVGGSGVVAOMSIKSNMGWAMSR 186
 DB 150 NKRVKXQKSGIRFPAISGHDYFELVYTVTNVGGSGVVAOMSIKSNMGWAMSR 209
 QY 187 QSNAYLAGOSLSFTVQDDGKRVYAMNXPXNM 219
 DB 210 QSNAYLAGOSLSFTVQDDGKRVYAMNXPXNM 242

RESULT 4
 Q9XHW9 PRELIMINARY: PRT: 254 AA.
 ID Q9XHW9
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE PUTATIVE EXPANSIN.
 GN T49B20.24 OR P0034C11.29.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Euphorbiaceae; Oryzaeae; Oryza.
 NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. NIPPONBARE;
 RA Buell C.R., Benito M.-I., Lin X., Mason T.M., Umayam L., Shea T.P.,
 RA Fujii C.Y., Shen M., Fraser C.M.;
 RT "Oryza sativa BAC T49B20 genomic sequence."
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
 clone:P0034C11."
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 11, 2002, 14:51:00 ; Search time 19.6007 Seconds
(Without alignments)
2003.488 Million cell updates/sec

Title: US-09-896-301-4

Perfect score: 1197

Sequence: 1 KXSVASQAFATFYGGKDGSC.....KVTAMNAPXNWLXXXXXX 227

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL.19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1173	98.0	255	10	Q40637 oryza sativ
2	885.5	74.0	251	10	Q9XHX0 oryza sativ
3	866.5	72.4	255	10	Q9FY29 festuca pra
4	809.5	67.6	254	10	Q9XHW9 oryza sativ
5	773.5	64.6	250	10	Q39625 cucumis sat
6	772.5	64.5	250	10	Q94619 oryza sativ
7	770.5	64.4	251	10	Q946J1 oryza sativ
8	769	64.2	256	10	Q92P39 nicotiana t
9	767	64.1	256	10	Q94KTS zea mays (m
10	765	63.9	257	10	Q9FVH0 lycopersico
11	764.5	63.9	252	10	Q91NU3 arabidopsis
12	762.5	63.7	251	10	Q40636 oryza sativ
13	760.5	63.5	252	10	Q9FUM2 prunus aviu
14	759.5	63.5	253	10	Q9FYJ1 festuca pra
15	759	63.4	252	10	Q9FY30 festuca pra
16	757.5	63.3	247	10	Q82625 lycopersico

17	756	63.2	261	10	Q24208	024208 oryza sativ
18	755.5	63.1	260	10	Q9FHT0	Q9FHT0 cicer ariet
19	751	62.7	256	10	Q9ZP38	Q9ZP38 nicotiana t
20	750.5	62.7	232	10	Q93493	Q93493 pinus taeda
21	750.5	62.7	232	10	Q93495	Q93495 pinus taeda
22	750.5	62.7	253	10	Q9SWY1	Q9SWY1 pinus taeda
23	750	62.7	257	10	Q9ZP37	Q9ZP37 nicotiana t
24	747.5	62.4	232	10	Q93492	Q93492 pinus taeda
25	747.5	62.4	254	10	Q82093	Q82093 prunus arme
26	746.5	62.4	254	10	Q9FUM3	Q9FUM3 prunus aviu
27	746.5	62.4	257	10	Q9XG16	Q9XG16 lycopersico
28	746.5	62.4	260	10	Q9M2S9	Q9M2S9 arabidopsis
29	744.5	62.2	247	10	Q9M5I7	Q9M5I7 triphsaria
30	744.5	62.2	249	10	Q91DR9	Q91DR9 arabidopsis
31	743.5	62.1	246	10	Q93442	Q93442 oryza sativ
32	743.5	62.1	249	10	Q9ZP35	Q9ZP35 nicotiana t
33	743.5	62.1	250	10	Q9FVG9	Q9FVG9 lycopersico
34	742.5	62.0	246	10	Q946J0	Q946J0 oryza sativ
35	742.5	62.0	264	10	Q9MAX6	Q9MAX6 oryza sativ
36	742.5	62.0	267	10	Q946H9	Q946H9 oryza sativ
37	740.5	61.9	245	10	Q91LB2	Q91LB2 zinnia eleg
38	740.5	61.9	253	10	Q94KT7	Q94KT7 zea mays (m
39	740	61.8	257	10	Q9FYS7	Q9FYS7 striga asia
40	738.5	61.7	249	10	Q9M5I5	Q9M5I5 triphsaria
41	738.5	61.7	260	10	Q81999	Q81999 lycopersico
42	737.5	61.6	232	10	Q93494	Q93494 pinus taeda
43	736.5	61.5	253	10	Q93XP2	Q93XP2 prunus cera
44	736.5	61.5	258	10	Q39626	Q39626 cucumis sat
45	736.5	61.5	262	10	Q946I7	Q946I7 oryza sativ

ALIGNMENTS

RESULT 1

Q40637 PRELIMINARY: PRT: 255 AA.
AC Q40637:
BT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
ID EXPANSIN OS-EXP3.
GN OS-EXP3.
OS Oryza sativa (Rice)
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIPPONBARE;
RX MEDLINE=96016146; PubMed=7568110;
RA Shcherban T.Y., Shi J., Durachko D.M., Gullittan M.J.,
RA McQueen-Mason S.J., Shieh M., Cosgrove D.J.;
RT "Molecular cloning and sequence analysis of expansins--a highly
RT conserved, multigene family of proteins that mediate cell wall
RT extension in plants.";
RT extension in plants.";
RT Proc. Natl. Acad. Sci. U.S.A. 92:9245-9249(1995).
RL EMBL: U30479; AAB38075.1;-
DR InterPro: IPR000882; Pollen_allergen.
DR Pfam: PF01357; Pollen_allergen; 1.
DR PRINTS: PR01225; EXPANSINFAMILY.
DR PRODOM: PD002179; Pollen_allergen; 1.
SQ SEQUENCE 255 AA: 27262 MW: 54AD66A196A1A648 CRC64;

Query Match

Best Local Similarity 98.0%; Score 1173; DB 10; Length 255;
Matches 214; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KXSVASQAFATFYGGKDGSCITMGACGYGNLYNAGYGLYNAALSSALFNDGACACCTT 60
DB 24 KXSVASQAFATFYGGKDGSCITMGACGYGNLYNAGYGLYNAALSSALFNDGACACCTT 83

PR	22-JUL-1999	99US-0145087
PR	22-JUL-1999	99US-0145088
PR	22-JUL-1999	99US-0145089
PR	22-JUL-1999	99US-0145192
PR	23-JUL-1999	99US-0145145
PR	23-JUL-1999	99US-0145148
PR	23-JUL-1999	99US-0145224
PR	26-JUL-1999	99US-0145226
PR	27-JUL-1999	99US-0145913
PR	27-JUL-1999	99US-0145914
PR	27-JUL-1999	99US-0145919
PR	28-JUL-1999	99US-0145951
PR	02-AUG-1999	99US-0146386
PR	02-AUG-1999	99US-0146388
PR	02-AUG-1999	99US-0146589
PR	03-AUG-1999	99US-0147038
PR	04-AUG-1999	99US-0147302
PR	04-AUG-1999	99US-0147304
PR	05-AUG-1999	99US-0147762
PR	05-AUG-1999	99US-0147760
PR	06-AUG-1999	99US-0147303
PR	06-AUG-1999	99US-0147416
PR	09-AUG-1999	99US-0147493
PR	09-AUG-1999	99US-0147935
PR	10-AUG-1999	99US-0148171
PR	11-AUG-1999	99US-0148319
PR	12-AUG-1999	99US-0148341
PR	13-AUG-1999	99US-0148565
PR	13-AUG-1999	99US-0148684
PR	16-AUG-1999	99US-0149368
PR	17-AUG-1999	99US-0149426
PR	18-AUG-1999	99US-0149475
PR	20-AUG-1999	99US-0149722
PR	20-AUG-1999	99US-0149923
PR	20-AUG-1999	99US-0149929
PR	23-AUG-1999	99US-0149902
PR	23-AUG-1999	99US-0149930
PR	25-AUG-1999	99US-0150566
PR	26-AUG-1999	99US-0150884
PR	27-AUG-1999	99US-0151065
PR	27-AUG-1999	99US-0151060
PR	27-AUG-1999	99US-0151080
PR	30-AUG-1999	99US-0151303
PR	31-AUG-1999	99US-0151438
PR	01-SEP-1999	99US-0151930
PR	07-SEP-1999	99US-0152363
PR	10-SEP-1999	99US-0153070
PR	13-SEP-1999	99US-0153758
PR	15-SEP-1999	99US-0154018
PR	16-SEP-1999	99US-0154039
PR	20-SEP-1999	99US-0154779
PR	22-SEP-1999	99US-0155139
PR	23-SEP-1999	99US-0155486
PR	24-SEP-1999	99US-0155659
PR	28-SEP-1999	99US-0156458
PR	29-SEP-1999	99US-0156596
PR	04-OCT-1999	99US-0157117
PR	05-OCT-1999	99US-0157753
PR	06-OCT-1999	99US-0157865
PR	07-OCT-1999	99US-0158029
PR	08-OCT-1999	99US-0158322
PR	13-OCT-1999	99US-0158369
PR	13-OCT-1999	99US-0159293
PR	13-OCT-1999	99US-0159294
PR	14-OCT-1999	99US-0159325
PR	14-OCT-1999	99US-0159330
PR	14-OCT-1999	99US-0159331
PR	14-OCT-1999	99US-0159637
PR	18-OCT-1999	99US-0159684
PR	21-OCT-1999	99US-0160741
PR	21-OCT-1999	99US-0160747
PR	21-OCT-1999	99US-0160768

PR	21-OCT-1999;	99US-01608174
PR	21-OCT-1999;	99US-01608110
PR	21-OCT-1999;	99US-01608115
PR	22-OCT-1999;	99US-01609880
PR	22-OCT-1999;	99US-01609881
PR	22-OCT-1999;	99US-01609889
PR	23-OCT-1999;	99US-01614005
PR	23-OCT-1999;	99US-01614004
PR	23-OCT-1999;	99US-01614059
PR	23-OCT-1999;	99US-01614056
PR	26-OCT-1999;	99US-01613609
PR	26-OCT-1999;	99US-01613660
PR	26-OCT-1999;	99US-01613621
PR	28-OCT-1999;	99US-01619932
PR	28-OCT-1999;	99US-01619923
PR	28-OCT-1999;	99US-01619933
PR	29-OCT-1999;	99US-01621442

Query Match	77.0%;	Score 927.5;	DB 21;	Length 259;
Best Local Similarity	73.6%;	Pred. No. 8.3e-95;		
Matches 167;	Conservative 22;	Mismatches 29;	Indels 9;	Gaps 2;

[illegible]

Search completed: October 11, 2002, 14:56:41
Job time : 24.5756 secs

PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match

Best Local Similarity 77.0%; Score 927.5; DB 21; Length 259;
Matches 167; Conservative 22; Mismatches 29; Indels 9; Gaps 2;

OY 3 GPMINAHATPEYXXGDXAXXTGACGAGNLYSQGYGLETALSTALPDGICGACXELMC 62
DB 34 GWMINAHATPEYGGCDASCTMGACGCGNLXSQGYGLETALSTALPDGICGACXELMC 93
OY 63 VNDPQICIGRSTIVTATNFCPP-----GGACDPNHHFDLSQPIYKIALYKSGIT 114
DB 94 ENDKKMCLEP-SIVVTATNFCPPNNALANNNGCMCPLEHFDLQAPVQRIAYRAGIV 152
OY 115 PVMYRVRCKRSGIRFTINGHSYFVLVTNCGAGDVHSYKMGSRKQWLMSSRMGO 174
DB 153 PVSIRVRCKRSGIRFTINGHSYFVLVTNCGAGDVHSYKMGSRKQWLMSSRMGO 212
OY 175 NMOSNYLNGQSLSFVVTSDRSVVSFVNAPPTWFSFGQTYTGQPR 221
DB 213 NMOSNYLNGQSLSFVVTSDRSVVSFVNAPPTWFSFGQTYTGQPR 259

RESULT 15

AGS1631
ID AGS1631 standard; Protein: 259 AA.

XX AGS1631;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 65549.

KM Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.

OS Arabidopsis thaliana.

PN EPI033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126284.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128234.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137528.
PR 07-JUN-1999; 99US-0137502.
PR 08-JUN-1999; 99US-0137724.
PR 10-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 28-JUN-1999; 99US-0140695.
PR 29-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0140991.
PR 01-JUL-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0142977.
PR 14-JUL-1999; 99US-0143542.
PR 15-JUL-1999; 99US-0143624.
PR 16-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.

PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0136021.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 18-JUN-1999; 99US-0139817.
PR 21-JUN-1999; 99US-0139899.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 23-JUN-1999; 99US-0140695.
PR 24-JUN-1999; 99US-0140823.
PR 28-JUN-1999; 99US-0140991.
PR 29-JUN-1999; 99US-0141287.
PR 30-JUN-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143342.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 19-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 20-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.

PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151308.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 13-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160960.
PR 22-OCT-1999; 99US-0160981.

```

PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147439.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151080.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159293.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.

```

```

PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Query Match 77.0%; Score 927.5; DB 21; Length 249;
 Best Local Similarity 73.6%; Pred. No. 7.8e-95;
 Matches 167; Conservative 22; Mismatches 29; Indels 9; Gaps 2;

```

QY 3 GPMINAHATFYXXGDAXXTWGGAGCGYGNLYSOGYGETALSTALSTALDOGLSCACKEIMK 62
DB 24 GCMINAHATFYGGDASGTWGGAGCGYGNLYSOGYGETALSTALSTALNNGLSCGCEIRK 83
QY 63 VNDPQWCIKRSIVYATNCP-----GGACDPNNHFDLSQPIYEKIALYKSGII 114
DB 84 ENDGKWCLPG-SIVYVATNFCPPNNMLANNNGCMCNPLEHFDLPVFORIAQYRAGIV 142
QY 115 PVMYRRYRCKRSGGIRPTINGHSYFNLVLTNTNGAGDVHSVSMKSRKQWLSRNMCO 174
DB 143 PVSTRYPCRRRCGIRPTINGHSYFNLVLTNTNGAGDVHSAIKSRVWQMSRNMCO 202
QY 175 NMQSNSTLNGQSLSFVYTTSDRSVVSFNVAPPTMSFGQTYTGQPR 221
DB 203 NMQSNSTLNGQSLSFVYTTSDGRTVVSFNAPACGNSYGTQFFAGQPR 249

```

RESULT 14
 AAG06544
 AAG06544 standard; Protein; 259 AA.

XX AAG06544;

DT 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 3355.

DE Protein identification: signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KM termination sequence.

XX Arabidopsis thaliana.

OS EPI033405-A2.

PD 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

```

XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0134256.

```


PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139873.
PR 22-JUN-1999; 99US-0139889.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140981.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.

PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159325.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.

PR	06-AUG-1999;	99US-0147416.	
PR	09-AUG-1999;	99US-0147493.	
PR	09-AUG-1999;	99US-0147935.	
PR	10-AUG-1999;	99US-0148171.	
PR	11-AUG-1999;	99US-0148319.	
PR	12-AUG-1999;	99US-0148341.	
PR	13-AUG-1999;	99US-0148565.	
PR	13-AUG-1999;	99US-0148684.	
PR	16-AUG-1999;	99US-0149368.	
PR	17-AUG-1999;	99US-0149175.	
PR	18-AUG-1999;	99US-0149426.	
PR	20-AUG-1999;	99US-0149722.	
PR	20-AUG-1999;	99US-0149723.	
PR	20-AUG-1999;	99US-0149929.	
PR	23-AUG-1999;	99US-0149902.	
PR	23-AUG-1999;	99US-0149930.	
PR	25-AUG-1999;	99US-0150566.	
PR	26-AUG-1999;	99US-0150884.	
PR	27-AUG-1999;	99US-0151065.	
PR	27-AUG-1999;	99US-0151066.	
PR	27-AUG-1999;	99US-0151080.	
PR	30-AUG-1999;	99US-0151303.	
PR	31-AUG-1999;	99US-0151438.	
PR	01-SEP-1999;	99US-0151930.	
PR	07-SEP-1999;	99US-0152363.	
PR	10-SEP-1999;	99US-0153070.	
PR	13-SEP-1999;	99US-0153758.	
PR	15-SEP-1999;	99US-0154018.	
PR	16-SEP-1999;	99US-0154039.	
PR	20-SEP-1999;	99US-0154779.	
PR	22-SEP-1999;	99US-0155139.	
PR	23-SEP-1999;	99US-0155486.	
PR	24-SEP-1999;	99US-0155659.	
PR	28-SEP-1999;	99US-0156458.	
PR	29-SEP-1999;	99US-0156596.	
PR	04-OCT-1999;	99US-0157117.	
PR	05-OCT-1999;	99US-0157753.	
PR	06-OCT-1999;	99US-0157865.	
PR	07-OCT-1999;	99US-0158029.	
PR	08-OCT-1999;	99US-0158232.	
PR	12-OCT-1999;	99US-0158369.	
PR	13-OCT-1999;	99US-0158293.	
PR	13-OCT-1999;	99US-0159294.	
PR	13-OCT-1999;	99US-0159295.	
PR	14-OCT-1999;	99US-0159329.	
PR	14-OCT-1999;	99US-0159330.	
PR	14-OCT-1999;	99US-0159331.	
PR	14-OCT-1999;	99US-0159637.	
PR	14-OCT-1999;	99US-0159638.	
PR	18-OCT-1999;	99US-0159584.	
PR	21-OCT-1999;	99US-0160741.	
PR	21-OCT-1999;	99US-0160767.	
PR	21-OCT-1999;	99US-0160776.	
PR	21-OCT-1999;	99US-0160778.	
PR	21-OCT-1999;	99US-0160814.	
PR	21-OCT-1999;	99US-0160815.	
PR	22-OCT-1999;	99US-0160980.	
PR	22-OCT-1999;	99US-0160981.	
PR	22-OCT-1999;	99US-0160989.	
PR	25-OCT-1999;	99US-0161404.	
PR	25-OCT-1999;	99US-0161405.	
PR	25-OCT-1999;	99US-0161406.	
PR	26-OCT-1999;	99US-0161359.	
PR	26-OCT-1999;	99US-0161360.	
PR	26-OCT-1999;	99US-0161361.	
PR	28-OCT-1999;	99US-0161920.	
PR	28-OCT-1999;	99US-0161992.	
PR	28-OCT-1999;	99US-0161993.	
PR	29-OCT-1999;	99US-0162142.	
Query Match			77.0%;
Best Local Similarity			73.6%;
Matches 167; Conservative			22;
			Score 927.5; DB 21;
			Length 249;
			Pred. No. 7.8e-95;
			Mismatches 29;
			Indels 9;
			Gaps 2

OY		3	GPMINAAHTEFEXXDGDAAAXMTWGAGGYNLVSQGGLETFALSTALFPOGLSCGACXELMC	62
D6		24	GWMIINAHAFTFFGGDGASGTWGCAGCYGNLSQGIGTSTAALSTRALFNNGLSGCSCEPIRC	83
OY		63	VNDPOMCTKRGSRIVVTATNFCPP-----GGACDPNNHHFDLSOPITYEKIALYRSGII	114
D6		84	ENDGMKCPLFG-SIVYTATNFPCPNNALANNNGWCNPPEHFHLDAQVFORIAQYRAGIV	142
OY		115	PVMTRVRVCCKRSGGIRFPFINSHSFNLVLVTNVGGADVDHSVSKSGRTTWOLMSRWMO	174
D6		143	PVSYSRRVPCRRRGGRFRFTINGHSYFNVLTLTNVGAGDVSHAALKGSRTYWAQMSRWMO	202
OY		175	NMOSNYLLNGOSSLFVVTTSDRRSVSPFNAPPTWSFGOTYTGQOF	221
D6		203	NMOSNYLLNGOALSFKYTTSDGRITVSPFNAPPAQMSICQTFFAGQOF	249
<hr/>				
RESULT 12				
AAG51632				
ID	AAG51632	standard; Protein; 249 AA.		
XX	AC	AAG51632;		
XX	DI	18-OCT-2000 (first entry)		
DE	Arabidopsis thaliana	protein fragment SEQ ID NO: 65550.		
XX	KW	Protein identification; signal transduction pathway; metabolic pathway;		
KW	hybridisation assay; genetic mapping; gene expression control; promoter	termination sequence.		
XX	OS	Arabidopsis thaliana.		
XX	PN	EPI033405-A2.		
XX	PD	06-SEP-2000.		
XX	PF	25-FEB-2000; 2000EP-0301439.		
XX	PR	25-FEB-1999; 99US-0121825.		
PR	05-MAR-1999;	99US-0123180.		
PR	09-MAR-1999;	99US-0123548.		
PR	23-MAR-1999;	99US-0125788.		
PR	25-MAR-1999;	99US-0126264.		
PR	29-MAR-1999;	99US-0126785.		
PR	01-APR-1999;	99US-0127462.		
PR	06-APR-1999;	99US-0128234.		
PR	08-APR-1999;	99US-0128714.		
PR	16-APR-1999;	99US-0129845.		
PR	19-APR-1999;	99US-0130077.		
PR	21-APR-1999;	99US-0130449.		
PR	23-APR-1999;	99US-0130510.		
PR	23-APR-1999;	99US-0130891.		
PR	28-APR-1999;	99US-0131444.		
PR	30-APR-1999;	99US-0132048.		
PR	30-APR-1999;	99US-0132407.		
PR	04-MAY-1999;	99US-0132484.		
PR	05-MAY-1999;	99US-0132485.		
PR	06-MAY-1999;	99US-0132486.		
PR	06-MAY-1999;	99US-0132487.		
PR	07-MAY-1999;	99US-0132863.		
PR	11-MAY-1999;	99US-0134256.		
PR	14-MAY-1999;	99US-0134218.		
PR	14-MAY-1999;	99US-0134213.		
PR	14-MAY-1999;	99US-0134221.		
PR	14-MAY-1999;	99US-0134370.		
PR	18-MAY-1999;	99US-0134768.		
PR	19-MAY-1999;	99US-0134941.		
PR	20-MAY-1999;	99US-0135124.		
PR	21-MAY-1999;	99US-0135353.		
PR	24-MAY-1999;	99US-0135629.		
PR	25-MAY-1999;	99US-0136021.		

PR 16-JUN-1999: 99US-0139452.
PR 16-JUN-1999: 99US-0139453.
PR 17-JUN-1999: 99US-0139452.
PR 18-JUN-1999: 99US-0139454.
PR 18-JUN-1999: 99US-0139455.
PR 18-JUN-1999: 99US-0139456.
PR 18-JUN-1999: 99US-0139457.
PR 18-JUN-1999: 99US-0139458.
PR 18-JUN-1999: 99US-0139459.
PR 18-JUN-1999: 99US-0139460.
PR 18-JUN-1999: 99US-0139461.
PR 18-JUN-1999: 99US-0139462.
PR 18-JUN-1999: 99US-0139463.
PR 18-JUN-1999: 99US-0139750.
PR 18-JUN-1999: 99US-0139763.
PR 21-JUN-1999: 99US-0139817.
PR 22-JUN-1999: 99US-0139899.
PR 23-JUN-1999: 99US-0140353.
PR 23-JUN-1999: 99US-0140354.
PR 24-JUN-1999: 99US-0140695.
PR 28-JUN-1999: 99US-0140823.
PR 29-JUN-1999: 99US-0140991.
PR 30-JUN-1999: 99US-0141287.
PR 01-JUL-1999: 99US-0141842.
PR 02-JUL-1999: 99US-0142154.
PR 02-JUL-1999: 99US-0142055.
PR 06-JUL-1999: 99US-0142390.
PR 08-JUL-1999: 99US-0142803.
PR 09-JUL-1999: 99US-0142920.
PR 12-JUL-1999: 99US-0142977.
PR 13-JUL-1999: 99US-0143562.
PR 14-JUL-1999: 99US-0143624.
PR 15-JUL-1999: 99US-0144005.
PR 16-JUL-1999: 99US-0144085.
PR 16-JUL-1999: 99US-0144086.
PR 19-JUL-1999: 99US-0144325.
PR 19-JUL-1999: 99US-0144331.
PR 19-JUL-1999: 99US-0144332.
PR 19-JUL-1999: 99US-0144333.
PR 19-JUL-1999: 99US-0144334.
PR 19-JUL-1999: 99US-0144335.
PR 20-JUL-1999: 99US-0144352.
PR 20-JUL-1999: 99US-0144632.
PR 20-JUL-1999: 99US-0144684.
PR 20-JUL-1999: 99US-0144884.
PR 21-JUL-1999: 99US-0144884.
PR 21-JUL-1999: 99US-0145086.
PR 21-JUL-1999: 99US-0145088.
PR 22-JUL-1999: 99US-0145085.
PR 22-JUL-1999: 99US-0145087.
PR 22-JUL-1999: 99US-0145089.
PR 22-JUL-1999: 99US-0145192.
PR 22-JUL-1999: 99US-0145192.
PR 23-JUL-1999: 99US-0145192.
PR 23-JUL-1999: 99US-0145218.
PR 23-JUL-1999: 99US-0145224.
PR 26-JUL-1999: 99US-0145276.
PR 27-JUL-1999: 99US-0145913.
PR 27-JUL-1999: 99US-0145918.
PR 27-JUL-1999: 99US-0145919.
PR 28-JUL-1999: 99US-0145951.
PR 02-AUG-1999: 99US-0146386.
PR 02-AUG-1999: 99US-0146388.
PR 02-AUG-1999: 99US-0146388.
PR 03-AUG-1999: 99US-0146388.
PR 03-AUG-1999: 99US-0147038.
PR 04-AUG-1999: 99US-0147204.
PR 04-AUG-1999: 99US-0147302.
PR 05-AUG-1999: 99US-0147192.
PR 05-AUG-1999: 99US-0147260.
PR 06-AUG-1999: 99US-0147303.
PR 06-AUG-1999: 99US-0147416.
PR 09-AUG-1999: 99US-0147453.
PR 09-AUG-1999: 99US-0147935.
PR 10-AUG-1999: 99US-0148171.
PR 11-AUG-1999: 99US-0148319.

PR 12-AUG-1999: 99US-0148341.
PR 13-AUG-1999: 99US-0148565.
PR 13-AUG-1999: 99US-0148684.
PR 16-AUG-1999: 99US-0149368.
PR 17-AUG-1999: 99US-0149175.
PR 18-AUG-1999: 99US-0149426.
PR 20-AUG-1999: 99US-0149722.
PR 20-AUG-1999: 99US-0149723.
PR 20-AUG-1999: 99US-0149929.
PR 23-AUG-1999: 99US-0149902.
PR 23-AUG-1999: 99US-0149930.
PR 25-AUG-1999: 99US-0150566.
PR 26-AUG-1999: 99US-0150884.
PR 27-AUG-1999: 99US-0151065.
PR 27-AUG-1999: 99US-0151066.
PR 27-AUG-1999: 99US-0151080.
PR 30-AUG-1999: 99US-0151303.
PR 31-AUG-1999: 99US-0151303.
PR 01-SEP-1999: 99US-0151930.
PR 07-SEP-1999: 99US-0152363.
PR 10-SEP-1999: 99US-0153070.
PR 13-SEP-1999: 99US-0153758.
PR 15-SEP-1999: 99US-0154018.
PR 16-SEP-1999: 99US-0154039.
PR 20-SEP-1999: 99US-0154779.
PR 22-SEP-1999: 99US-0155139.
PR 23-SEP-1999: 99US-0155486.
PR 24-SEP-1999: 99US-0155659.
PR 28-SEP-1999: 99US-0156458.
PR 29-SEP-1999: 99US-0156596.
PR 04-OCT-1999: 99US-0157117.
PR 05-OCT-1999: 99US-0157753.
PR 06-OCT-1999: 99US-0157865.
PR 07-OCT-1999: 99US-0158029.
PR 08-OCT-1999: 99US-0158232.
PR 12-OCT-1999: 99US-0158369.
PR 13-OCT-1999: 99US-0159293.
PR 13-OCT-1999: 99US-0159294.
PR 13-OCT-1999: 99US-0159295.
PR 14-OCT-1999: 99US-0159329.
PR 14-OCT-1999: 99US-0159330.
PR 14-OCT-1999: 99US-0159331.
PR 14-OCT-1999: 99US-0159637.
PR 14-OCT-1999: 99US-0159638.
PR 18-OCT-1999: 99US-0159584.
PR 21-OCT-1999: 99US-0160741.
PR 21-OCT-1999: 99US-0160767.
PR 21-OCT-1999: 99US-0160768.
PR 21-OCT-1999: 99US-0160770.
PR 21-OCT-1999: 99US-0160814.
PR 21-OCT-1999: 99US-0160815.
PR 22-OCT-1999: 99US-0160980.
PR 22-OCT-1999: 99US-0160981.
PR 22-OCT-1999: 99US-0160989.
PR 25-OCT-1999: 99US-0161404.
PR 25-OCT-1999: 99US-0161405.
PR 25-OCT-1999: 99US-0161406.
PR 26-OCT-1999: 99US-0161359.
PR 26-OCT-1999: 99US-0161360.
PR 26-OCT-1999: 99US-0161361.
PR 26-OCT-1999: 99US-0161361.
PR 28-OCT-1999: 99US-0161920.
PR 28-OCT-1999: 99US-0161992.
PR 28-OCT-1999: 99US-0161993.
PR 29-OCT-1999: 99US-0162142.

Query Match 77.0%; Score 927.5; DB 21; Length 241;
Best Local Similarity 73.6%; Pred. No. 7, 5e-95;
Matches 167; Conservative 22; Mismatches 29; Indels 9; Gaps 2;

OY 3 GPTNHAATFYXXGDAAXXTMGACAGYGNIXSOGYLETALSTALPFDGIGSCACXELMC 62
DB 16 GWTNHAATFYXGGDASGTMGACAGYGNIXSOGYGTALSTALPNNNGISCSGCFEINC 75

```
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154772.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
```

Query Match 77.08; Score 927.5; DB 21; Length 241;
Best Local Similarity 73.68; Pred. No. 7.5e-95;
Matches 167; Conservative 22; Mismatches 29; Indels 9; Gaps 2;

```
OY 3 GPMINAHATFYXXDAXXKMGAGCYGNLYSOGYGLTALSTALPFOGLSCGACXKLMC 62
DB 16 GCMINAHATFYXXDAXXKMGAGCYGNLYSOGYGLTALSTALPFOGLSCGACXKLMC 75
OY 63 VNPOMKIKGRSIVYATNFCRP-----GCACDPNNHFDLSQPIYEKIALYKSGII 114
DB 76 ENDKMKLPG-SIVVYATNFCRPNNALANNNGWGNPRLHFRLDQVFGRIQYRAGIV 134
OY 115 PVMYRVRVRCRSGIRPTINGHSYFNLYLVITNNGAGADIVASVSKSGRTKWLMSRNGO 174
```

```
DB 135 PVSYRVRVRCRSGIRPTINGHSYFNLYLVITNNGAGADIVASVSKSGRTKWLMSRNGO 194
OY 175 NMOSNYLNGQSLFSFVYTTSDRSVSEFNVAAPPYMSFGQYTGQFR 221
DB 195 NMOSNYLNGQSLFSFVYTTSDRSVSEFNVAAPPYMSFGQYTGQFR 241
```

RESULT 10

AAGS1647 standard; Protein: 241 AA.

AAGS1647;

18-OCT-2000 (first entry)

Arabidopsis thaliana protein fragment SEQ ID NO: 65570.

KW Protein identification: signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PE 25-FEB-2000; 2000EP-0301439.

```
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130891.
XX 30-APR-1999; 99US-0131449.
XX 04-MAY-1999; 99US-0132407.
XX 05-MAY-1999; 99US-0132484.
XX 06-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 07-MAY-1999; 99US-0132487.
XX 11-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134376.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
```



```
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139887.
PR 22-JUN-1999; 99US-0139889.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 28-JUN-1999; 99US-0140695.
PR 29-JUN-1999; 99US-0140883.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141882.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0142977.
PR 14-JUL-1999; 99US-0143542.
PR 15-JUL-1999; 99US-0143624.
PR 16-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145313.
PR 27-JUL-1999; 99US-0145319.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 12-AUG-1999; 99US-0148319.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149829.
PR 23-AUG-1999; 99US-0149902.

PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155139.
PR 24-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0155659.
PR 29-SEP-1999; 99US-0156458.
PR 05-OCT-1999; 99US-0156596.
PR 05-OCT-1999; 99US-0157117.
PR 06-OCT-1999; 99US-0157753.
PR 07-OCT-1999; 99US-0157863.
PR 08-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161929.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 77.0%; Score 927.5; DB 21; Length 241;
Best Local Similarity 73.6%; Pred. No. 7.5e-95;
Matches 167; Conservative 22; Mismatches 29; Indels 9; Gaps 2;
```

PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155655.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158025.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159684.
PR 18-OCT-1999; 99US-0159588.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
Query Match 91.0%; Score 1096; DB 21; Length 201;
Best Local Similarity 99.5%; Pred. No. 1e-113;
Matches 200; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 22 MGAGCGYGNLYSGYGTETALSTALPFDGLSCGACXELKCVNDPWCICGRSIVYATN 81
DB 1 MGAGCGYGNLYSGYGTETALSTALPFDGLSCGACXELKCVNDPWCICGRSIVYATN 60
QY 82 FCPGACDPNNHFDLSQPIYERIALYKSGIIPVYRVRCKRSGSIRFTINGHSYENL 141
DB 61 FCPGACDPNNHFDLSQPIYERIALYKSGIIPVYRVRCKRSGSIRFTINGHSYENL 120
QY 142 VLVTVNGAGDVHSVSKSGRTKQWLSRWGQWQNSNLTNGOSLSFVYTTSDRSYVS 201
DB 121 VLVTVNGAGDVHSVSKSGRTKQWLSRWGQWQNSNLTNGOSLSFVYTTSDRSYVS 180
QY 202 FNVAPPWPSFGQYTYGQFPRY 222
DB 181 FNVAPPWPSFGQYTYGQFPRY 201
RESULT 8
AAG06546

ID AAG06546 standard; Protein: 241 AA.
XX AAG06546;
AC AAG06546;
DT 17-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 3357.
DE Arabidopsis thaliana protein fragment SEQ ID NO: 3357.
XX Protein identification: signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
PM EPI033405-A2.
PD 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
PF 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 14-MAY-1999; 99US-0134376.
PR 18-MAY-1999; 99US-0134944.
PR 19-MAY-1999; 99US-0135124.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139453.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 64704.
DE Protein identification: signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 23-MAR-1999; 99US-0126284.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131448.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134236.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.

PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.

```

PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 03-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145951.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.

```

```

PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Query Match 91.0%; Score 1096; DB 21; Length 201;
 Best Local Similarity 99.5%; Pred. No. 1e-113;
 Matches 200; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 22 MGGACGYGNLYSOGYLETALSTALFDGLSCGACXELMCVNDPOMCTIKRSIVVTATN 81
    |||||||
DB 1 MGGACGYGNLYSOGYLETALSTALFDGLSCGACFELMCVNDPOMCTIKRSIVVTATN 60
    |||||||
QY 82 FCPPGACDPPNHHFDSLQPIYERKIALYSGIIPVYRRVCKRSGIRFTINGHSYFNL 141
    |||||||
DB 61 FCPPGACDPPNHHFDSLQPIYERKIALYSGIIPVYRRVCKRSGIRFTINGHSYFNL 120
    |||||||
QY 142 VLVTNWGGAGDVHSVSMKSGRTKQWOLMSRWGNQMSNYLNGQSLFVVTSDRSYVS 201
    |||||||
DB 121 VLVTNWGGAGDVHSVSMKSGRTKQWOLMSRWGNQMSNYLNGQSLFVVTSDRSYVS 180
    |||||||
QY 202 FNVAPPTWSFGQTYTGQGFY 222
    |||||||
DB 181 FNVAPPTWSFGQTYTGQGFY 201
    |||||||

```

RESULT 7
 ID AAG51014
 AC AAG51014;
 XX
 XX 18-OCT-2000 (first entry)

KM	termination sequence.
XX	
OS	<i>Arabidopsis thaliana</i> .
PN	EP1033405-A2.
XX	
PD	06-SEP-2000.
XX	
XX	25-FEB-2000; 2000EP-0301439
XX	
PR	25-FEB-1999; 9905-01218235
PR	05-MAR-1999; 9905-0123160
PR	09-MAR-1999; 9905-0123548
PR	23-MAR-1999; 9905-0125768
PR	23-MAR-1999; 9905-0126264
PR	23-MAR-1999; 9905-0126785
PR	01-APR-1999; 9905-0127462
PR	08-APR-1999; 9905-0128234
PR	08-APR-1999; 9905-0128714
PR	16-APR-1999; 9905-0129845
PR	13-APR-1999; 9905-0130077
PR	21-APR-1999; 9905-0130449
PR	23-APR-1999; 9905-0130510
PR	23-APR-1999; 9905-0130891
PR	28-APR-1999; 9905-0131449
PR	30-APR-1999; 9905-0132048
PR	30-APR-1999; 9905-0132407
PR	04-MAY-1999; 9905-0132484
PR	04-MAY-1999; 9905-0132485
PR	06-MAY-1999; 9905-0132486
PR	06-MAY-1999; 9905-0132487
PR	07-MAY-1999; 9905-0132863
PR	11-MAY-1999; 9905-0134256
PR	14-MAY-1999; 9905-0134218
PR	14-MAY-1999; 9905-0134221
PR	14-MAY-1999; 9905-0134221
PR	14-MAY-1999; 9905-0134370
PR	18-MAY-1999; 9905-0134768
PR	19-MAY-1999; 9905-0134941

PR	25-MAY-1999:	99US-0136021
PR	24-MAY-1999:	99US-0135629
PR	21-MAY-1999:	99US-0135353

PR	18-JUN-1999	9905-0136532
PR	28-MAY-1999	9905-0136782
PR	18-JUN-1999	9905-0136788
PR	03-JUN-1999	9905-0137222
PR	03-JUN-1999	9905-0137528
PR	07-JUN-1999	9905-0137502
PR	07-JUN-1999	9905-0137724
PR	08-JUN-1999	9905-0138024
PR	10-JUN-1999	9905-0138340
PR	10-JUN-1999	9905-0138847
PR	14-JUN-1999	9905-0139152
PR	16-JUN-1999	9905-0139452
PR	16-JUN-1999	9905-0139653
PR	17-JUN-1999	9905-0139692
PR	18-JUN-1999	9905-0139654
PR	18-JUN-1999	9905-0139655
PR	18-JUN-1999	9905-0139656

PR	18-JUN-1999;	9905-0139450
PR	18-JUN-1999;	9905-0139460
PR	18-JUN-1999;	9905-0139461
PR	18-JUN-1999;	9905-0139466
PR	18-JUN-1999;	9905-0139467
PR	18-JUN-1999;	9905-0139468
PR	18-JUN-1999;	9905-0139469
PR	18-JUN-1999;	9905-0139750
PR	21-JUN-1999;	9905-0139763
PR	21-JUN-1999;	9905-0139817
PR	22-JUN-1999;	9905-0139869
PR	23-JUN-1999;	9905-0140353
PR	23-JUN-1999;	9905-0140354
PR	24-JUN-1999;	9905-0140695

XX 06-SEP-2000.
PD 2000EP-0301439.
XX 25-FEB-2000;
XX 25-FEB-1999;
PR 05-MAR-1999;
PR 09-MAR-1999;
PR 23-MAR-1999;
PR 25-MAR-1999;
PR 01-APR-1999;
PR 06-APR-1999;
PR 16-APR-1999;
PR 19-APR-1999;
PR 21-APR-1999;
PR 23-APR-1999;
PR 28-APR-1999;
PR 30-APR-1999;
PR 04-MAY-1999;
PR 05-MAY-1999;
PR 06-MAY-1999;
PR 07-MAY-1999;
PR 11-MAY-1999;
PR 14-MAY-1999;
PR 14-MAY-1999;
PR 14-MAY-1999;
PR 18-MAY-1999;
PR 19-MAY-1999;
PR 20-MAY-1999;
PR 21-MAY-1999;
PR 24-MAY-1999;
PR 25-MAY-1999;
PR 27-MAY-1999;
PR 28-MAY-1999;
PR 01-JUN-1999;
PR 03-JUN-1999;
PR 04-JUN-1999;
PR 07-JUN-1999;
PR 08-JUN-1999;
PR 10-JUN-1999;
PR 10-JUN-1999;
PR 14-JUN-1999;
PR 16-JUN-1999;
PR 17-JUN-1999;
PR 18-JUN-1999;
PR 18-JUN-1999;
PR 18-JUN-1999;
PR 18-JUN-1999;
PR 18-JUN-1999;
PR 21-JUN-1999;
PR 22-JUN-1999;
PR 23-JUN-1999;
PR 23-JUN-1999;
PR 24-JUN-1999;
PR 28-JUN-1999;
PR 29-JUN-1999;
PR 30-JUN-1999;
PR 01-JUL-1999;
PR 01-JUL-1999;

99US-0121825.
99US-0123180.
99US-0123548.
99US-0125788.
99US-0126264.
99US-0126785.
99US-0127462.
99US-0128234.
99US-0128714.
99US-0129845.
99US-0130077.
99US-0130449.
99US-0130510.
99US-0130891.
99US-0131449.
99US-0132048.
99US-0132407.
99US-0132484.
99US-0132485.
99US-0132486.
99US-0132487.
99US-0132863.
99US-0134256.
99US-0134219.
99US-0134221.
99US-0134370.
99US-0134768.
99US-0134941.
99US-0135124.
99US-0135353.
99US-0135629.
99US-0136021.
99US-0136392.
99US-0136782.
99US-0137222.
99US-0137528.
99US-0137502.
99US-0137724.
99US-0138094.
99US-0138447.
99US-0138847.
99US-0139119.
99US-0139452.
99US-0139453.
99US-0139492.
99US-0139454.
99US-0139455.
99US-0139456.
99US-0139457.
99US-0139458.
99US-0139459.
99US-0139460.
99US-0139461.
99US-0139462.
99US-0139463.
99US-0139750.
99US-0139753.
99US-0139817.
99US-0139899.
99US-0140353.
99US-0140354.
99US-0140655.
99US-0140823.
99US-0140991.
99US-0141287.
99US-0141842.
99US-0142154.

PR 02-JUL-1999;
PR 06-JUL-1999;
PR 08-JUL-1999;
PR 09-JUL-1999;
PR 12-JUL-1999;
PR 13-JUL-1999;
PR 14-JUL-1999;
PR 15-JUL-1999;
PR 16-JUL-1999;
PR 16-JUL-1999;
PR 19-JUL-1999;
PR 19-JUL-1999;
PR 19-JUL-1999;
PR 19-JUL-1999;
PR 19-JUL-1999;
PR 20-JUL-1999;
PR 21-JUL-1999;
PR 21-JUL-1999;
PR 21-JUL-1999;
PR 22-JUL-1999;
PR 22-JUL-1999;
PR 22-JUL-1999;
PR 22-JUL-1999;
PR 23-JUL-1999;
PR 23-JUL-1999;
PR 23-JUL-1999;
PR 26-JUL-1999;
PR 27-JUL-1999;
PR 27-JUL-1999;
PR 27-JUL-1999;
PR 28-JUL-1999;
PR 02-AUG-1999;
PR 02-AUG-1999;
PR 03-AUG-1999;
PR 04-AUG-1999;
PR 04-AUG-1999;
PR 05-AUG-1999;
PR 05-AUG-1999;
PR 06-AUG-1999;
PR 06-AUG-1999;
PR 09-AUG-1999;
PR 09-AUG-1999;
PR 10-AUG-1999;
PR 11-AUG-1999;
PR 12-AUG-1999;
PR 13-AUG-1999;
PR 13-AUG-1999;
PR 16-AUG-1999;
PR 17-AUG-1999;
PR 18-AUG-1999;
PR 20-AUG-1999;
PR 20-AUG-1999;
PR 20-AUG-1999;
PR 23-AUG-1999;
PR 23-AUG-1999;
PR 25-AUG-1999;
PR 26-AUG-1999;
PR 27-AUG-1999;
PR 27-AUG-1999;
PR 30-AUG-1999;
PR 31-AUG-1999;
PR 01-SEP-1999;
PR 07-SEP-1999;
PR 10-SEP-1999;
PR 13-SEP-1999;
PR 15-SEP-1999;
PR 16-SEP-1999;
PR 20-SEP-1999;

99US-0142055.
99US-0142390.
99US-0142803.
99US-0142920.
99US-0142977.
99US-0143542.
99US-0143624.
99US-0144005.
99US-0144085.
99US-0144086.
99US-0144325.
99US-0144331.
99US-0144332.
99US-0144333.
99US-0144334.
99US-0144335.
99US-0144352.
99US-0144632.
99US-0144884.
99US-0144814.
99US-0145086.
99US-0145088.
99US-0145085.
99US-0145087.
99US-0145089.
99US-0145192.
99US-0145145.
99US-0145218.
99US-0145224.
99US-0145276.
99US-0145913.
99US-0145918.
99US-0145919.
99US-0145951.
99US-0146386.
99US-0146388.
99US-0146389.
99US-0147038.
99US-0147204.
99US-0147302.
99US-0147192.
99US-0147260.
99US-0147303.
99US-0147416.
99US-0147493.
99US-0147935.
99US-0148171.
99US-0148319.
99US-0148341.
99US-0148565.
99US-0148684.
99US-0149368.
99US-0149175.
99US-0149426.
99US-0149722.
99US-0149723.
99US-0149929.
99US-0149902.
99US-0149930.
99US-0150566.
99US-0150884.
99US-0151065.
99US-0151066.
99US-0151080.
99US-0151303.
99US-0151348.
99US-0151930.
99US-0152363.
99US-0153070.
99US-0153758.
99US-0154018.
99US-0154039.
99US-0154779.

```

PR 13-JUL-1999; 99US-013542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145192.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147058.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147416.
PR 06-AUG-1999; 99US-0147433.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.

```

```

PR 04-OCT-1999; 99US-0157171.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158332.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159337.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Query Match 98.6%; Score 1187; DB 21; Length 221;
Best Local Similarity 97.7%; Pred. No. 8, 5e-124;
Matches 216; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 2 MGPWINAHTEFYXXGDAAXXTMGACGCGYGNLYSGCYLETAALSTALFPDGLSCGACXELM 61
DB 1 MGPWINAHTEFYXXGDAAXXTMGACGCGYGNLYSGCYLETAALSTALFPDGLSCGACXELM 60
QY 62 CVNDPQMCIKGRSIVVTATNFCPPGACDPPNHHFDSLQPIYEKIALYKSGIIPVYRRV 121
DB 61 CVNDPQMCIKGRSIVVTATNFCPPGACDPPNHHFDSLQPIYEKIALYKSGIIPVYRRV 120
QY 122 RCKRSGCIRPTINGHSYFNLVLTNNGAGDVHSYMKGSRTKWLQMSRNMCMQNSY 181
DB 121 RCKRSGCIRPTINGHSYFNLVLTNNGAGDVHSYMKGSRTKWLQMSRNMCMQNSY 180
QY 182 LNCOSLSFVYTTSDRSYVSFNVAPPTMSFGOTYTGOPRY 222
DB 181 LNCOSLSFVYTTSDRSYVSFNVAPPTMSFGOTYTGOPRY 221

```

```

RESULT 5
AAG51013
ID AAG51013 standard; Protein; 221 AA.
XX
AC AAG51013;
XX
DT 18-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment seq ID NO: 64703.
DE
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.

```


PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159638.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160880.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 99.3%; Score 1195; DB 21; Length 255;
Best Local Similarity 97.7%; Pred. No. 1.3e-124;
Matches 217; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 HMGFWINAHATFYXXGDAXXTMGACGYGMLYSOGYGLETAALSTALFDGSLSCGACXEL 60
|||||
DB 34 HMGFWINAHATFYGGDASCTMGACGYGMLYSOGYGLETAALSTALFDGSLSCGACFEL 93
|||||
OY 61 MCVNDPOMCIKGRSIYVATNFPCPGACDPPNHHFDLSQPIETKIALTKSGIIPMYRR 120
|||||
DB 94 MCVNDPOMCIKGRSIYVATNFPCPGACDPPNHHFDLSQPIETKIALTKSGIIPMYRR 153
|||||
OY 121 VRCKRSGRIFETINGHYENLVLTWVGAGDVHSYMKGSRTKQWLMSPNQMSNS 180
|||||
DB 154 VRCKRSGRIFETINGHYENLVLTWVGAGDVHSYMKGSRTKQWLMSPNQMSNS 213
|||||
OY 181 YLNGQSLSFVYVTTSDRSVSVFNVAAPPTWSFGQTYTGQFRY 222
|||||
DB 214 YLNGQSLSFVYVTTSDRSVSVFNVAAPPTWSFGQTYTGQFRY 255
|||||

RESULT 2

ID AAG51012 standard; Protein; 255 AA.

AC AAG51012;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 64702.

XX Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139765.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.

PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134257.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135253.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139819.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 28-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0143624.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.

PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150565.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

```
Run on:      October 11, 2002, 14:49:55 ; Search time 24.5756 Seconds
              (without alignments)
              1003.367 Million cell updates/sec
```

Title: US-09-896-301-3

Sequence: 1 HMGPTNAHATFYXXGDAXX.....NVAPPTWSFGQTYTGQFRY 222

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing:	Minimum Match	0%
	Maximum Match	100%
	Listing first	45 summaries

A.Geneseq.032802.*

1:	/SDSI/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
2:	/SDSI/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
3:	/SDSI/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
4:	/SDSI/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
5:	/SDSI/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
6:	/SDSI/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
7:	/SDSI/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
8:	/SDSI/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
9:	/SDSI/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
10:	/SDSI/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
11:	/SDSI/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
12:	/SDSI/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
13:	/SDSI/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
14:	/SDSI/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
15:	/SDSI/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
16:	/SDSI/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
17:	/SDSI/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
18:	/SDSI/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
19:	/SDSI/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
20:	/SDSI/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
21:	/SDSI/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
22:	/SDSI/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1195	99.3	255	21	AA615693	Arabidopsis thalia
2	1195	99.3	255	21	AA615692	Arabidopsis thalia
3	1194	99.2	222	17	AA694529	Rice expansin. Or
4	1187	98.6	221	21	AA615694	Arabidopsis thalia
5	1187	98.6	221	21	AA651013	Arabidopsis thalia
6	1096	91.0	201	21	AA615695	Arabidopsis thalia
7	1096	91.0	201	21	AA651014	Arabidopsis thalia
8	927.5	77.0	241	21	AA606446	Arabidopsis thalia
9	927.5	77.0	241	21	AA651633	Arabidopsis thalia
10	927.5	77.0	241	21	AA651647	Arabidopsis thalia
11	927.5	77.0	249	21	AA606545	Arabidopsis thalia

12	927.5	77.0	249	21	AA651632	Arabidopsis thaliana
13	927.5	77.0	259	21	AA651646	Arabidopsis thaliana
14	927.5	77.0	259	21	AA605544	Arabidopsis thaliana
15	927.5	77.0	289	21	AA651631	Arabidopsis thaliana
16	927.5	77.0	280	21	AA651645	Arabidopsis thaliana
17	907.5	75.4	251	21	AA623852	Arabidopsis thaliana
18	907.5	75.4	251	21	AA643343	Arabidopsis thaliana
19	907.5	75.4	253	21	AA623851	Arabidopsis thaliana
20	907.5	75.4	253	21	AA643342	Arabidopsis thaliana
21	907.5	75.4	281	21	AA643341	Arabidopsis thaliana
22	907.5	75.4	282	21	AA623850	Arabidopsis thaliana
23	989.5	74.6	228	17	AA694528	Rice expansin. OR
24	876.5	72.8	241	21	AA605453	Arabidopsis thaliana
25	876.5	72.8	249	21	AA605452	Arabidopsis thaliana
26	876.5	72.8	255	21	AA605451	Arabidopsis thaliana
27	871.5	72.4	242	21	AA636570	Arabidopsis thaliana
28	871.5	72.4	249	21	AA636569	Arabidopsis thaliana
29	871.5	72.4	255	21	AA636568	Arabidopsis thaliana
30	857.5	71.2	227	17	AA694527	Cucumber expansin-7
31	855.5	71.1	262	21	AA629931	Arabidopsis thaliana
32	855.5	71.1	273	21	AA629930	Arabidopsis thaliana
33	853.5	70.9	250	22	AA600414	Tomato seed expansin
34	845.5	70.2	250	21	AA609622	Arabidopsis thaliana
35	845.5	70.2	258	21	AA609621	Arabidopsis thaliana
36	845.5	70.2	280	21	AA609620	Arabidopsis thaliana
37	842	69.9	253	21	AA625443	Arabidopsis thaliana
38	842	69.9	253	21	AA64483	Arabidopsis thaliana
39	832.5	69.1	257	21	AA636445	Arabidopsis thaliana
40	804.5	66.8	255	21	AA630325	Arabidopsis thaliana
41	804.5	66.8	257	21	AA630324	Arabidopsis thaliana
42	796.5	66.2	263	22	AA600412	Tomato seed expansin
43	790	65.6	207	21	AA625444	Arabidopsis thaliana
44	790	65.6	207	21	AA646844	Arabidopsis thaliana
45	782.5	65.0	210	21	AA629932	Arabidopsis thaliana

RESULT 1

AA615693

AA615693 standard; Protein; 255 AA.

XX	AA615693:	
AC	AA615693:	
XX	17-OCT-2000 (first entry)	
DT		
XX	Arabidopsis thaliana protein fragment SEQ ID NO: 16047.	
DE		
XX	Protein identification: signal transduction pathway; metabolic pathway;	
KW	hybridisation assay; genetic mapping; gene expression control; promoter;	
KW	termination sequence.	
XX		
XX	Arabidopsis thaliana.	
OS		
NN	EP1033405-A2.	
XX		
PD	06-SEP-2000.	
XX		
PF	25-FEB-2000; 2000EP-0301439.	
XX		
PR	25-FEB-1999; 99US-0121825.	
PR	05-MAR-1999; 99US-0123180.	
PR	09-MAR-1999; 99US-0123548.	
PR	23-MAR-1999; 99US-0125788.	
PR	25-MAR-1999; 99US-0126264.	
PR	29-MAR-1999; 99US-0126785.	
PR	01-APR-1999; 99US-0127462.	
PR	06-APR-1999; 99US-0128234.	
PR	08-APR-1999; 99US-0128714.	
PR	16-APR-1999; 99US-0129845.	
PR	19-APR-1999; 99US-0130077.	
PR	21-APR-1999; 99US-0130449.	

ALIGNMENTS

[illegible]


```

? ? ? ? ?
? CURRENT FILING DATE: 1998-05-01
? EARLIER APPLICATION NUMBER: 60/045,445
? EARLIER FILING DATE: 1997-05-02
? NUMBER OF SEQ ID NOS: 37
? SOFTWARE: PatentIn Ver. 2.1
? SEQ ID NO 37
? LENGTH: 254
? TYPE: PRT
? ORGANSISM: Arabidopsis
US-09-071-252-37
```

Query Match	75.5%	Score 909.5;	DB 14;	Length 254;
Best Local Similarity	73.1%	Pred. No. 8.2e-89;		
Matches 166;	Conservative 19;	Mismatches 33;	Indels 9;	Gaps 2;

Qy	3	GPATINAHATFFXXKDDAXXTHMGACGYCNLSQOAGLEFALSTLTPQOGLSCGACXELMC	62
Db	29	GGWYNAHAHTFFGGDDASGTHMGACGYCNLSQOCTGTALSTLFPNGLSCGACFEIRC	88
Qy	63	VNDPQWCIKGRHSIVTATNFPCP-----GGACDPPNHHFDSLQPIYERIALYKSGII	114
Db	89	QNDGKCWCLPG-SIVVTATNFCPPNNALPNNAGCNCNPPQGHFDSLQGVFQRIHQYRAGIV	147
Qy	115	PVMTRRVRRCRSGGIRFTINGHSYFNLLVLTNNGACDVSHSVSKSGRTYQWOLMSRWGO	174
Db	148	PVAARVPVCVRGGIRFTINGHSYFNLLVLTNNGACDVSHSVSKSGRTYQWOLMSRWGO	207
Qy	175	NMQNSYTLNGOSLSFVVTTSRBRVSVFNPAPTWSSGORYTYGOQR	221
Db	208	NMQNSYTLNGOSLSFVKVTTSGOYIVSNVNAAMGWSGGOFPTGQOLR	254

```

RESULT 15
US-09-708-427-35961
: Sequence 35961, Application US/09708427
: GENERAL INFORMATION:
: APPLICANT: N. ALEXANDROV et al.
: TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
: TITLE OF INVENTION: THEREBY
: FILE REFERENCE: 2750-1243P
: CURRENT APPLICATION NUMBER: US/09/708,427
: CURRENT FILING DATE: 2000-11-09
: NUMBER OF SEQ ID NOS: 85364
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 35961
: LENGTH: 235
: TYPE: PRT
: ORGANISM: Arabidopsis thaliana
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 1..235
: OTHER INFORMATION: Xaa is any amino acid
: NAME/KEY: misc_feature
: LOCATION: 1..235
: OTHER INFORMATION: Ceres Seq. ID 1842533
: US-09-708-427-35961

```

[illegible]

DB 39 GGWTAHATFYGGSDASGIMGGACGYNLYSQGYGTNTAALSTALFNNGJSCGSCYEIRC 98

sequence 11, application US/03071222C
GENERAL INFORMATION:
APPLICANT: Cosgrove, Daniel J
TITLE OF INVENTION: BETA-EXPANSINS AS CELL WALL LOOSENING AGENTS

1 TITLE OF INVENTION: COMPOSITIONS THEREOF AND METHODS OF USE
2
3
4 FILE REFERENCE: 11940E183
5
6 CURRENT APPLICATION NUMBER: US/09/071.252C

LENGTH: 249
TYPE: PRT
ORGANISM: Arabidopsis thaliana

OY	3	GPMVINAATYTYXXGDXAXXMXMGACGTCYMLYSQGTGLEETALSTALPDGLSCGACXELMC	62
Db	24	GCWVINAATYTYGGGDXSGTMSGACGTCYMLYSQGTSTALSTALENNGLSCGSCFEIJC	83
OY	63	VNDFQWCIKRSIVYATNFPCPP-----GGACDPENHIEPLSPDIYEKIALYKKSIT	114
Db	84	ENDGKMCIDP-SIVYATNFPCPPNALALANNNGGCMNPLEHEDLPAPFQRIAOYRAIV	142
OY	115	PWYARVRCCKRSGGIRFPIINGHSIFNLVLTYNVGAGDVHSYMSGRTKQMLSRMNGO	174
Db	143	PVSTRNRCRRGGIRFPIINGHSYFNLVLTYNVGAGDVHSAIIGSRVQWQMSRMNGO	202
OY	175	NMQSNSYLGQSLPVTYSDRSRVSFNVAAPPTAFSGQTYTGQPR	221
Db	203	NMQSNSYLGQSLPVTYSDGRTVSFNAAPGMSYGGTFAFGGQPR	249

RESULT 11
US-09-595-329A-2001
; Sequence 2001, Application US/09595329A

```

Query Match          99.2%; Score 1194; DB 22; Length 222;
Best Local Similarity 100.0%; Pred. No. 1,8e-119;
Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HMGPMINAHATFYXXGDAXXTMGACGYGNLYSOGYGLETAALSTALFDQGLSCGACXEL 60
    |||
DB 1 HMGPMINAHATFYXXGDAXXTMGACGYGNLYSOGYGLETAALSTALFDQGLSCGACXEL 60
OY 61 MCNDPQWCIKGRSIVVTATNFCPPGACDPNNHFDLSQPIYEKIALYKSGIIPMYRR 120
    |||
DB 61 MCNDPQWCIKGRSIVVTATNFCPPGACDPNNHFDLSQPIYEKIALYKSGIIPMYRR 120
OY 121 VRCKRSGGIRFTINGHSYFNLYLVNMGAGDVHSYMSKSRKQWOLMSRMNGQWQNS 180
    |||
DB 121 VRCKRSGGIRFTINGHSYFNLYLVNMGAGDVHSYMSKSRKQWOLMSRMNGQWQNS 180
OY 121 VRCKRSGGIRFTINGHSYFNLYLVNMGAGDVHSYMSKSRKQWOLMSRMNGQWQNS 180
    |||
DB 121 VRCKRSGGIRFTINGHSYFNLYLVNMGAGDVHSYMSKSRKQWOLMSRMNGQWQNS 180
OY 181 YLNGOSLSFVYTTSDRRSVSFVNAPPTWSFGQTYTGQFRY 222
    |||
DB 181 YLNGOSLSFVYTTSDRRSVSFVNAPPTWSFGQTYTGQFRY 222

RESULT 5
US-09-620-394B-4043
; Sequence 4043, Application US/09620394B
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai
; APPLICANT: BROVER, Vyacheslav
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 2750-1067P
; CURRENT APPLICATION NUMBER: US/09/620,394B
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 9131
; SEQ ID NO 4043
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..221
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: 1..221
; OTHER INFORMATION: Ceres Seq. ID 1388406
US-09-620-394B-4043

Query Match          98.6%; Score 1187; DB 20; Length 221;
Best Local Similarity 97.7%; Pred. No. 9,9e-119;
Matches 216; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 2 MGFMINAHATFYXXGDAXXTMGACGYGNLYSOGYGLETAALSTALFDQGLSCGACXEL 61
    |||
DB 1 MGFMINAHATFYXXGDAXXTMGACGYGNLYSOGYGLETAALSTALFDQGLSCGACXEL 60
OY 62 CVNDPQWCIKGRSIVVTATNFCPPGACDPNNHFDLSQPIYEKIALYKSGIIPMYRR 121
    |||
DB 61 CVNDPQWCIKGRSIVVTATNFCPPGACDPNNHFDLSQPIYEKIALYKSGIIPMYRR 120
OY 122 RCKRSGGIRFTINGHSYFNLYLVNMGAGDVHSYMSKSRKQWOLMSRMNGQWQNS 181
    |||
DB 121 RCKRSGGIRFTINGHSYFNLYLVNMGAGDVHSYMSKSRKQWOLMSRMNGQWQNS 180
OY 121 RCKRSGGIRFTINGHSYFNLYLVNMGAGDVHSYMSKSRKQWOLMSRMNGQWQNS 180
    |||
DB 121 RCKRSGGIRFTINGHSYFNLYLVNMGAGDVHSYMSKSRKQWOLMSRMNGQWQNS 180
OY 182 YLNGOSLSFVYTTSDRRSVSFVNAPPTWSFGQTYTGQFRY 222
    |||
DB 181 YLNGOSLSFVYTTSDRRSVSFVNAPPTWSFGQTYTGQFRY 221

RESULT 6
US-09-620-394B-4044
; Sequence 4044, Application US/09620394B
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai
; APPLICANT: BROVER, Vyacheslav

```

```

; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 2750-1067P
; CURRENT APPLICATION NUMBER: US/09/620,394B
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 9131
; SEQ ID NO 4044
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..201
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: 1..201
; OTHER INFORMATION: Ceres Seq. ID 1388407
US-09-620-394B-4044

Query Match          91.0%; Score 1096; DB 20; Length 201;
Best Local Similarity 99.5%; Pred. No. 5,4e-109;
Matches 200; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 22 MGCAGYGNLYSOGYGLETAALSTALFDQGLSCGACXELMCVNDPQWCIKGRSIVVTATN 81
    |||
DB 1 MGCAGYGNLYSOGYGLETAALSTALFDQGLSCGACXELMCVNDPQWCIKGRSIVVTATN 60
OY 82 FCPREGACDPNNHFDLSQPIYEKIALYKSGIIPMYRRVRCKRSGGIRFTINGHSYFN 141
    |||
DB 61 FCPREGACDPNNHFDLSQPIYEKIALYKSGIIPMYRRVRCKRSGGIRFTINGHSYFN 120
OY 142 VLVTVNGAGDVHSYMSKSRKQWOLMSRMNGQWQNSYLNQOSLSFVYTTSDRRSVS 201
    |||
DB 121 VLVTVNGAGDVHSYMSKSRKQWOLMSRMNGQWQNSYLNQOSLSFVYTTSDRRSVS 180
OY 202 FNVAPPTWSFGQTYTGQFRY 222
    |||
DB 181 FNVAPPTWSFGQTYTGQFRY 201

RESULT 7
US-60-324-109-21420
; Sequence 21420, Application US/60324109
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Edgerly, Michael D
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jindong
; APPLICANT: Stein, Joshua
; TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-10(52726)B
; CURRENT APPLICATION NUMBER: US/60/324,109
; CURRENT FILING DATE: 2001-09-21
; NUMBER OF SEQ ID NOS: 33196
; SEQ ID NO 21420
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: Glycine max
US-60-324-109-21420

Query Match          83.4%; Score 1004.5; DB 26; Length 247;
Best Local Similarity 80.0%; Pred. No. 4,8e-99;
Matches 176; Conservative 22; Mismatches 21; Indels 1; Gaps 1;

OY 3 GPMINAHATFYXXGDAXXTMGACGYGNLYSOGYGLETAALSTALFDQGLSCGACXEL 62
    |||
DB 29 GPMINAHATFYXXGDAXXTMGACGYGNLYSOGYGLETAALSTALFDQGLSCGACXEL 88
OY 63 VNDPQWCIKGRSIVVTATNFCPPGACDPNNHFDLSQPIYEKIALYKSGIIPMYRR 122
    |||
DB 89 VNDPQWCIKGRSIVVTATNFCPPGACDPNNHFDLSQPIYEKIALYKSGIIPMYRR 147

```

```

Db 12 HMGPMINAHATFYGGDASGTMGACGYGNLYSGYGLFETALSTALFDQGLSCGACFEL 71
QY 61 MCVNDPQMCIKGRSIVVTATNFCPPGACDPPNHHFDLSOPITYEKIALYKSGIIPVWYR 120
Db 72 MCVNDPQMCIKGRSIVVTATNFCPPGACDPPNHHFDLSOPITYEKIALYKSGIIPVWYR 131
QY 121 VRCKRSGGIRFTINGHSYFNLYLVTVNGAGADVHSVSMKGSRTKQWQLMSRNQGNQNS 180
Db 132 VRCKRSGGIRFTINGHSYFNLYLVTVNGAGADVHSVSMKGSRTKQWQLMSRNQGNQNS 191
QY 181 YLNGQSLSFVVTSDRRSVSFVNAPPTWSFGQTYTGQFRY 222
Db 192 YLNGQSLSFVVTSDRRSVSFVNAPPTWSFGQTYTGQFRY 233

```

RESULT 2

```

US-09-112-498A-9
Sequence 9, Application US/09112498A
GENERAL INFORMATION:

```

```

TITLE OF INVENTION: MICROBIAL SMOLLENIN PROTEIN, DNA SEQUENCES
TITLE OF INVENTION: ENCODING SUCH SMOLLENINS AND METHOD OF PRODUCING SUCH
NUMBER OF SEQUENCES: 31
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (ERO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/112,498A
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 98/14226
FILING DATE:
APPLICATION NUMBER: US 08/893,766
FILING DATE: 11-JUL-1997
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 233 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-09-112-498A-9

```

```

Query Match 99.3%: Score 1195; DB 15; Length 233;
Best Local Similarity 97.7%: Pred. No. 1.5e-119;
Matches 217; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 1 HMGPMINAHATFYXXGDAAXXTMGACGYGNLYSGYGLFETALSTALFDQGLSCGACFEL 60
Db 12 HMGPMINAHATFYGGDASGTMGACGYGNLYSGYGLFETALSTALFDQGLSCGACFEL 71
QY 61 MCVNDPQMCIKGRSIVVTATNFCPPGACDPPNHHFDLSOPITYEKIALYKSGIIPVWYR 120
Db 72 MCVNDPQMCIKGRSIVVTATNFCPPGACDPPNHHFDLSOPITYEKIALYKSGIIPVWYR 131
QY 121 VRCKRSGGIRFTINGHSYFNLYLVTVNGAGADVHSVSMKGSRTKQWQLMSRNQGNQNS 180
Db 132 VRCKRSGGIRFTINGHSYFNLYLVTVNGAGADVHSVSMKGSRTKQWQLMSRNQGNQNS 191
QY 181 YLNGQSLSFVVTSDRRSVSFVNAPPTWSFGQTYTGQFRY 222
Db 192 YLNGQSLSFVVTSDRRSVSFVNAPPTWSFGQTYTGQFRY 233

```

RESULT 3

```

US-09-620-394B-4042
Sequence 4042, Application US/09620394B
GENERAL INFORMATION:

```

```

APPLICANT: ALEXANDROV, Nickolai
APPLICANT: BROVER, Vyacheslav
TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides

```

```

TITLE OF INVENTION: Thereby
FILE REFERENCE: 2750-1067P
CURRENT APPLICATION NUMBER: US/09/620,394B
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 9131
SEQ ID NO 4042
LENGTH: 255
TYPE: PRT
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..255
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: 1..255
OTHER INFORMATION: Ceres Seq. ID 1388405
US-09-620-394B-4042

```

```

Query Match 99.3%: Score 1195; DB 20; Length 255;
Best Local Similarity 97.7%: Pred. No. 1.6e-119;
Matches 217; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 1 HMGPMINAHATFYXXGDAAXXTMGACGYGNLYSGYGLFETALSTALFDQGLSCGACFEL 60
Db 34 HMGPMINAHATFYGGDASGTMGACGYGNLYSGYGLFETALSTALFDQGLSCGACFEL 93
QY 61 MCVNDPQMCIKGRSIVVTATNFCPPGACDPPNHHFDLSOPITYEKIALYKSGIIPVWYR 120
Db 94 MCVNDPQMCIKGRSIVVTATNFCPPGACDPPNHHFDLSOPITYEKIALYKSGIIPVWYR 131
QY 121 VRCKRSGGIRFTINGHSYFNLYLVTVNGAGADVHSVSMKGSRTKQWQLMSRNQGNQNS 180
Db 154 VRCKRSGGIRFTINGHSYFNLYLVTVNGAGADVHSVSMKGSRTKQWQLMSRNQGNQNS 213
QY 181 YLNGQSLSFVVTSDRRSVSFVNAPPTWSFGQTYTGQFRY 222
Db 214 YLNGQSLSFVVTSDRRSVSFVNAPPTWSFGQTYTGQFRY 255

```

RESULT 4

```

US-09-896-301-3
Sequence 3, Application US/09896301
GENERAL INFORMATION:

```

```

APPLICANT: Cosgrove, Daniel J
APPLICANT: McQueen-Mason, Simon
APPLICANT: Gullitnan, Mark J
APPLICANT: Shcherban, Tatyana
TITLE OF INVENTION: PURIFIED EXPANSIN PROTEINS
FILE REFERENCE: 1194/1C114US3
CURRENT APPLICATION NUMBER: US/09/896,301
CURRENT FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: 09/092,160
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 08/440,517
PRIOR FILING DATE: 1995-05-12
PRIOR APPLICATION NUMBER: 08/242,090
PRIOR FILING DATE: 1994-05-12
PRIOR APPLICATION NUMBER: 08/060,944
PRIOR FILING DATE: 1993-05-12
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 222
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: rice expansin
NAME/KEY: UNSURE
LOCATION: (14)..(58)
OTHER INFORMATION: Xaa is unknown or other.
US-09-896-301-3

```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 11, 2002, 14:56:50 ; Search time 98.6303 Seconds
(without alignments)
792.249 Million cell updates/sec

Title: US-09-896-301-3

Perfect score: 1204
Sequence: 1 HMGWIMAHATFYXXGXDAAXX.....NVAPPTWGFQITTCGCTRY 222

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 3502263 seqs, 351980561 residues

Total number of hits satisfying chosen parameters: 3502263

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents, AA, Main:*

1: /cgn2_6/ptodata/2/paa/US086_COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US087_COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US088_COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US089_COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US090_COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US091_COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US092_COMB.pep.*
8: /cgn2_6/ptodata/2/paa/US093_COMB.pep.*
9: /cgn2_6/ptodata/2/paa/US094_COMB.pep.*
10: /cgn2_6/ptodata/2/paa/US095_COMB.pep.*
11: /cgn2_6/ptodata/2/paa/US096_COMB.pep.*
12: /cgn2_6/ptodata/2/paa/US097_COMB.pep.*
13: /cgn2_6/ptodata/2/paa/US098_COMB.pep.*
14: /cgn2_6/ptodata/2/paa/US099_COMB.pep.*
15: /cgn2_6/ptodata/2/paa/US100_COMB.pep.*
16: /cgn2_6/ptodata/2/paa/US101_COMB.pep.*
17: /cgn2_6/ptodata/2/paa/US102_COMB.pep.*
18: /cgn2_6/ptodata/2/paa/US103_COMB.pep.*
19: /cgn2_6/ptodata/2/paa/US104_COMB.pep.*
20: /cgn2_6/ptodata/2/paa/US105_COMB.pep.*
21: /cgn2_6/ptodata/2/paa/US106_COMB.pep.*
22: /cgn2_6/ptodata/2/paa/US107_COMB.pep.*
23: /cgn2_6/ptodata/2/paa/US108_COMB.pep.*
24: /cgn2_6/ptodata/2/paa/US109_COMB.pep.*
25: /cgn2_6/ptodata/2/paa/US110_COMB.pep.*
26: /cgn2_6/ptodata/2/paa/US111_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1195	99.3	233	US-09-112-498-10	Sequence 10, Appl
2	1195	99.3	233	US-09-112-498-9	Sequence 9, Appl
3	1195	99.3	255	US-09-620-394B-4042	Sequence 4042, Ap
4	1194	99.2	222	US-09-896-301-3	Sequence 3, Appl
5	1187	98.6	221	US-09-620-394B-4043	Sequence 4043, Ap
6	1096	91.0	201	US-09-620-394B-4044	Sequence 4044, Ap
7	1004.5	83.4	247	US-60-324-109-21420	Sequence 21420, A

8	927.5	77.0	242	US-09-595-329A-2003	Sequence 2003, Ap
9	927.5	77.0	249	US-09-570-581A-1823	Sequence 1823, Ap
10	927.5	77.0	250	US-09-595-329A-2002	Sequence 2002, Ap
11	927.5	77.0	260	US-09-595-329A-2001	Sequence 2001, Ap
12	913.5	75.9	264	US-60-324-109-21386	Sequence 21386, A
13	909.5	75.5	237	US-09-112-498A-7	Sequence 7, Appl
14	909.5	75.5	254	US-09-071-252-37	Sequence 35961, A
15	907.5	75.4	235	US-09-708-427-35961	Sequence 35961, A
16	907.5	75.4	248	US-09-708-427-35960	Sequence 35960, A
17	905.5	75.2	248	US-60-312-544-6608	Sequence 6608, Ap
18	904.5	75.1	236	US-09-595-329A-955	Sequence 955, App
19	904.5	75.1	250	US-09-595-329A-954	Sequence 954, App
20	898.5	74.6	228	US-09-896-301-2	Sequence 2, Appl
21	898	74.6	236	US-09-112-498-4	Sequence 8, Appl
22	891.5	74.0	223	US-09-112-498A-3	Sequence 3, Appl
23	891.5	74.0	223	US-09-595-328C-1914	Sequence 1914, Ap
24	876.5	72.8	242	US-09-595-328C-1913	Sequence 1913, Ap
25	876.5	72.8	250	US-09-595-328C-1912	Sequence 1912, Ap
26	876.5	72.8	256	US-09-112-498-5	Sequence 5, Appl
27	870	72.3	222	US-60-324-109-22664	Sequence 22664, A
28	868.5	72.1	227	US-09-112-498A-5	Sequence 5, Appl
29	857.5	71.2	227	US-09-896-301-7	Sequence 7, Appl
30	857.5	71.2	227	US-09-708-427-1757	Sequence 1757, Ap
31	855.5	71.1	262	US-09-708-427-1756	Sequence 1756, Ap
32	855.5	71.1	273	US-09-410-191-6	Sequence 6, Appl
33	853.5	70.9	250	US-09-071-252-23	Sequence 23, Appl
34	853.5	70.9	250	US-09-071-252-22	Sequence 22, Appl
35	848.5	70.5	250	US-09-689-960-152	Sequence 152, App
36	842	69.9	253	US-09-708-427-35962	Sequence 35962, A
37	841.5	69.9	207	US-60-324-109-22976	Sequence 22976, A
38	838.5	69.6	273	US-09-112-498-14	Sequence 14, Appl
39	832.5	69.1	237	US-09-708-427-2083	Sequence 2083, Ap
40	832.5	69.1	257	US-09-112-498A-13	Sequence 13, Appl
41	832.5	69.1	258	US-09-595-329A-956	Sequence 956, App
42	828.5	68.8	207	US-09-112-498-6	Sequence 6, Appl
43	827.5	68.7	225	US-09-112-498-8	Sequence 8, Appl
44	823	68.4	232		
45			15		

ALIGNMENTS

RESULT 1
US-09-112-498-10
; Sequence 10, Application US/09112498
; GENERAL INFORMATION:
; APPLICANT: Swanson, Barbara A.
; APPLICANT: Ward, Michael
; APPLICANT: Penttila, Merja
; APPLICANT: Jaakko, Pere
; APPLICANT: Saloheimo, Markku
; TITLE OF INVENTION: Microleins and Method of Producing Such Swollenins
; FILE REFERENCE: GC378-2
; CURRENT APPLICATION NUMBER: US/09/112,498
; CURRENT FILING DATE: 1998-07-09
; PRIOR APPLICATION NUMBER: 08/893,766
; PRIOR FILING DATE: 1997-07-11
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Plant/eukaryotic
US-09-112-498-10

Query Match 99.3%; Score 1195; DB 15; Length 233;
Best Local Similarity 97.7%; Pred. No. 1.5e-119;
Matches 217; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
1 HMGWIMAHATFYXXGXDAAXXGTGACGYGMLYSGYGLTFAALSTLFDGSCGACXEL 60
|||||

OY 115 PVMYRRVCKRSGIRFTINGSHYENLVLTNMGAGDVHVSXKGSRTKWQLMSRNMGO 174
115 PVMYRRVCKRSGIRFTINGSHYENLVLTNMGAGDVHVSXKGSRTKWQLMSRNMGO 174
Db 144 PVAYRRVPCVRRGIRFTINGSHYENLVLTNMGAGDVHVSXKGSRTKWQLMSRNMGO 203
144 PVAYRRVPCVRRGIRFTINGSHYENLVLTNMGAGDVHVSXKGSRTKWQLMSRNMGO 203
OY 175 NMOSNSYILNGSLSFVVTTSRDRSVSENVAPPTWSFGQTYTGQFR 221
175 NMOSNSYILNGSLSFVVTTSRDRSVSENVAPPTWSFGQTYTGQFR 221
Db 204 NMOSNSYILNGSLSFVVTTSRDRSVSENVAPPTWSFGQTYTGQFR 250
204 NMOSNSYILNGSLSFVVTTSRDRSVSENVAPPTWSFGQTYTGQFR 250

Search completed: October 11, 2002, 15:15:44
Job time : 31.293 secs

```
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: 1..250
; OTHER INFORMATION: Ceres Seq. ID no. 3446951
US-09-935-625-23149
```

```
Query Match          75.5%; Score 909.5; DB 5; Length 250;
Best Local Similarity 73.1%; Pred. No. 7.6e-88;
Matches 166; Conservative 19; Mismatches 33; Indels 9; Gaps 2;
```

```
QY 3 GPMINAHATFYXXGDAAXXTMGACGYGNLYSGYGLTALSTALFDDGLSCGACXELMC 62
   ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
Db 25 GGVVNAHATFYGGDASGTMGACGYGNLYSGYGTNTALSTALFNNGLSGACFEIRC 84
   ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
QY 63 VNDPQWCIKGRSIVTATNFCPP-----GGACDPNHHFDLSQPIYEKIALYKSGII 114
   ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
Db 85 QNDGKWCILPG-STIVTATNFCPPNNALPNNAGGWCNPPQOHFDLSQPVFGRIAGYRAGIV 143
   ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
QY 115 PWAYRRVRCRSGGIRFTINGHSYFNLVLTNNGAGDVHSVMKSGSRKWKQMLSRNMQ 174
   ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
Db 144 PWAYRRVPCVRKRGIRFTINGHSYFNLVLTNNGAGDVHSAMVKSGSRKWKQMLSRNMQ 203
   ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
QY 175 NMOSNSYILNGQSLSFVYTTSDRRSVSFNVAAPTWSFGOTYTGGOFR 221
   ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
Db 204 NMOSNSYILNGQSLSFVYTTSDGOTIVSNVANAAGWSFGOTFTGAQLR 250
   ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
```

RESULT 13
US-09-935-625-23153

```
; Sequence 23153, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPABLE
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 23153
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: 1..250
; OTHER INFORMATION: Ceres Seq. ID no. 3446955
US-09-935-625-23153
```

```
Query Match          75.5%; Score 909.5; DB 5; Length 250;
Best Local Similarity 73.1%; Pred. No. 7.6e-88;
Matches 166; Conservative 19; Mismatches 33; Indels 9; Gaps 2;
```

```
QY 3 GPMINAHATFYXXGDAAXXTMGACGYGNLYSGYGLTALSTALFDDGLSCGACXELMC 62
   ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
Db 25 GGVVNAHATFYGGDASGTMGACGYGNLYSGYGTNTALSTALFNNGLSGACFEIRC 84
   ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
QY 63 VNDPQWCIKGRSIVTATNFCPP-----GGACDPNHHFDLSQPIYEKIALYKSGII 114
   ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
Db 85 QNDGKWCILPG-STIVTATNFCPPNNALPNNAGGWCNPPQOHFDLSQPVFGRIAGYRAGIV 143
   ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
QY 115 PWAYRRVRCRSGGIRFTINGHSYFNLVLTNNGAGDVHSVMKSGSRKWKQMLSRNMQ 174
   ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
Db 144 PWAYRRVPCVRKRGIRFTINGHSYFNLVLTNNGAGDVHSAMVKSGSRKWKQMLSRNMQ 203
   ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
QY 175 NMOSNSYILNGQSLSFVYTTSDRRSVSFNVAAPTWSFGOTYTGGOFR 221
   ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
Db 204 NMOSNSYILNGQSLSFVYTTSDGOTIVSNVANAAGWSFGOTFTGAQLR 250
   ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
```

RESULT 14
US-09-935-625-31515

```
; Sequence 31515, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPA
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 31515
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: 1..250
; OTHER INFORMATION: Ceres Seq. ID no. 3446951
US-09-935-625-31515
```

```
Query Match          75.5%; Score 909.5; DB 5; Length 250;
Best Local Similarity 73.1%; Pred. No. 7.6e-88;
Matches 166; Conservative 19; Mismatches 33; Indels 9; Gaps 2;
```

```
QY 3 GPMINAHATFYXXGDAAXXTMGACGYGNLYSGYGLTALSTALFDDGLSCGACXELMC 62
   ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
Db 25 GGVVNAHATFYGGDASGTMGACGYGNLYSGYGTNTALSTALFNNGLSGACFEIRC 84
   ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
QY 63 VNDPQWCIKGRSIVTATNFCPP-----GGACDPNHHFDLSQPIYEKIALYKSGII 114
   ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
Db 85 QNDGKWCILPG-STIVTATNFCPPNNALPNNAGGWCNPPQOHFDLSQPVFGRIAGYRAGIV 143
   ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
QY 115 PWAYRRVRCRSGGIRFTINGHSYFNLVLTNNGAGDVHSVMKSGSRKWKQMLSRNMQ 174
   ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
Db 144 PWAYRRVPCVRKRGIRFTINGHSYFNLVLTNNGAGDVHSAMVKSGSRKWKQMLSRNMQ 203
   ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
QY 175 NMOSNSYILNGQSLSFVYTTSDRRSVSFNVAAPTWSFGOTYTGGOFR 221
   ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
Db 204 NMOSNSYILNGQSLSFVYTTSDGOTIVSNVANAAGWSFGOTFTGAQLR 250
   ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
```

RESULT 15
US-09-935-625-31519

```
; Sequence 31519, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPA
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 31519
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: 1..250
; OTHER INFORMATION: Ceres Seq. ID no. 3446955
US-09-935-625-31519
```

```
Query Match          75.5%; Score 909.5; DB 5; Length 250;
Best Local Similarity 73.1%; Pred. No. 7.6e-88;
Matches 166; Conservative 19; Mismatches 33; Indels 9; Gaps 2;
```

```
QY 3 GPMINAHATFYXXGDAAXXTMGACGYGNLYSGYGLTALSTALFDDGLSCGACXELMC 62
   ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
Db 25 GGVVNAHATFYGGDASGTMGACGYGNLYSGYGTNTALSTALFNNGLSGACFEIRC 84
   ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
QY 63 VNDPQWCIKGRSIVTATNFCPP-----GGACDPNHHFDLSQPIYEKIALYKSGII 114
   ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
Db 85 QNDGKWCILPG-STIVTATNFCPPNNALPNNAGGWCNPPQOHFDLSQPVFGRIAGYRAGIV 143
   ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
```


Db 130 PVATRRVPCVRRGIRFTINGHSYFNVLITVNWGAGDVHSAMVKGSRFGQTFTGAQLR 189
QY 175 NMOSNSTYLNQSLSFVYVTTSDRRSVSFVNAPPTWSPFGQTYTGQGR 221
Db 190 NMOSNSTYLNQSLSFVYVTTSDGQTIIVSNVANNAGMSFGQTFGTGAQLR 236

RESULT 9

US-09-935-625-31520
; Sequence 31520, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPABLE
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 31520
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: 1..236
; OTHER INFORMATION: Ceres Seq. ID no. 3446956
US-09-935-625-31520

Query Match 75.5%; Score 909.5; DB 5; Length 236;
Best Local Similarity 73.1%; Pred. No. 7.1e-88;
Matches 166; Conservative 19; Mismatches 33; Indels 9; Gaps 2;

QY 3 GPVINAHAATFYXXGDAAXXTMGACGYNLYSGYGLFETALSTALFDOGLSCGACXELMC 62
Db 11 GGVVNAHAATFYGGDASGTMGACGYNLYSGYGTNTALSTALFNNGLSGACFEIRC 70
QY 63 VNDPQWCIKGRSIVYVATNFCPP-----GCACDPNNHFDLSQPIYEKIALYKSGIT 114
Db 71 QNDGKWCPLPG-SIVYVATNFCPPNNALPNNAGMCNPPQOHFDLSQPIYFQIRAOYRAGIV 129
QY 115 PVMYRRVRCRKGIRFTINGHSYFNVLITVNWGAGDVHSAMVKGSRFGQTYTGQGR 174
Db 130 PVATRRVPCVRRGIRFTINGHSYFNVLITVNWGAGDVHSAMVKGSRFGQTYTGQGR 189
QY 175 NMOSNSTYLNQSLSFVYVTTSDRRSVSFVNAPPTWSPFGQTYTGQGR 221
Db 190 NMOSNSTYLNQSLSFVYVTTSDGQTIIVSNVANNAGMSFGQTFGTGAQLR 236

RESULT 10

US-09-935-625-2249
; Sequence 2249, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPABLE
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 2249
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: 1..250
; OTHER INFORMATION: Ceres Seq. ID no. 3446951
US-09-935-625-2249

Query Match 75.5%; Score 909.5; DB 5; Length 250;
Best Local Similarity 73.1%; Pred. No. 7.6e-88;
Matches 166; Conservative 19; Mismatches 33; Indels 9; Gaps 2;

QY 3 GPVINAHAATFYXXGDAAXXTMGACGYNLYSGYGLFETALSTALFDOGLSCGACXELMC 62
Db 25 GGVVNAHAATFYGGDASGTMGACGYNLYSGYGTNTALSTALFNNGLSGACFEIRC 84
QY 63 VNDPQWCIKGRSIVYVATNFCPP-----GCACDPNNHFDLSQPIYEKIALYKSGIT 114
Db 85 QNDGKWCPLPG-SIVYVATNFCPPNNALPNNAGMCNPPQOHFDLSQPIYFQIRAOYRAGIV 143
QY 115 PVMYRRVRCRKGIRFTINGHSYFNVLITVNWGAGDVHSAMVKGSRFGQTYTGQGR 174
Db 144 PVATRRVPCVRRGIRFTINGHSYFNVLITVNWGAGDVHSAMVKGSRFGQTYTGQGR 203
QY 175 NMOSNSTYLNQSLSFVYVTTSDRRSVSFVNAPPTWSPFGQTYTGQGR 221
Db 204 NMOSNSTYLNQSLSFVYVTTSDGQTIIVSNVANNAGMSFGQTFGTGAQLR 250

RESULT 11

US-09-935-625-2253
; Sequence 2253, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPABLE
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 2253
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: 1..250
; OTHER INFORMATION: Ceres Seq. ID no. 3446955
US-09-935-625-2253

Query Match 75.5%; Score 909.5; DB 5; Length 250;
Best Local Similarity 73.1%; Pred. No. 7.6e-88;
Matches 166; Conservative 19; Mismatches 33; Indels 9; Gaps 2;

QY 3 GPVINAHAATFYXXGDAAXXTMGACGYNLYSGYGLFETALSTALFDOGLSCGACXELMC 62
Db 25 GGVVNAHAATFYGGDASGTMGACGYNLYSGYGTNTALSTALFNNGLSGACFEIRC 84
QY 63 VNDPQWCIKGRSIVYVATNFCPP-----GCACDPNNHFDLSQPIYEKIALYKSGIT 114
Db 85 QNDGKWCPLPG-SIVYVATNFCPPNNALPNNAGMCNPPQOHFDLSQPIYFQIRAOYRAGIV 143
QY 115 PVMYRRVRCRKGIRFTINGHSYFNVLITVNWGAGDVHSAMVKGSRFGQTYTGQGR 174
Db 144 PVATRRVPCVRRGIRFTINGHSYFNVLITVNWGAGDVHSAMVKGSRFGQTYTGQGR 203
QY 175 NMOSNSTYLNQSLSFVYVTTSDRRSVSFVNAPPTWSPFGQTYTGQGR 221
Db 204 NMOSNSTYLNQSLSFVYVTTSDGQTIIVSNVANNAGMSFGQTFGTGAQLR 250

RESULT 12

US-09-935-625-23149
; Sequence 23149, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPABLE
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 23149
; LENGTH: 250

QY 3 GPRINNAHFEYXXGDAAXXIMGACGCGYGNLSQGVGLEFALSTALFPOGSLSCGACXKILMC 62

Db 11 GGVNNAHAFFYGGDASGCMGACGCGYGNLSQGVGTNAAALSTALFNNGLSCGACFEIRC 70

QY 63 VNDPQMCIGKRSIVTAINECP-----GGACDPNNHFFDLISQPIYKELIAYKSGII 114

Db 71 QNDGKRCCLPG-STVYATATNFCPPNNAALPNAGGWCNPPQOAFDLSQPEFQKIAIYRGIY 129

QY 115 PWVYRVRRCRKSGLIFETTINGSHYENLYVTWVGAGCVHSAVSMKSGRTKQMLSRMWQ 174

```

: GENERAL INFORMATION:
: APPLICANT: Cao, Yongwei
: APPLICANT: Edgerton, Michael D
: APPLICANT: Hinkle, Gregory J.
: APPLICANT: Kovalic, David K.
: APPLICANT: Liu, Jingdong
: APPLICANT: Stein, Joshua
: TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
: FILE REFERENCE: 38-10(52726)C
: CURRENT APPLICATION NUMBER: US/10/219,999
: PRIOR FILING DATE: 2002-08-15
: PRIOR APPLICATION NUMBER: US 60/324,109
: PRIOR FILING DATE: 2001-09-21
: PRIOR APPLICATION NUMBER: US 60/312,544
: PRIOR FILING DATE: 2001-08-15
: NUMBER OF SEQ ID NOS: 63520
: SEQ ID NO 59261
: LENGTH: 205
: TYPE: PRT
: ORGANISM: Glycine max
: US-10-219-999-59261

```

```

Query Match 79.1%; Score 952.5; DB 6; Length 205;
Best Local Similarity 81.1%; Pred. No. 1.7e-92;
Matches 167; Conservative 22; Mismatches 16; Indels 1; Gaps 1;

```

```

QY 17 DAXXTGACGAGYGNLYSGYGLTAAALSTALFDOGLSCGACXELMCVNDPOMCIRKRSIV 76
  |||||||
Db 1 DASGTGACGAGYGNLYSGYGTNTALSTALFNNGLSCGSCYQICANDPQCLMG-TIV 59
  |||||||
QY 77 VPATNCPGACDPNNHFDLSOPITKIALYKSGIIPVVRVRCMKRSGIRFTINGH 136
  |||||||
Db 60 VVATNCPGACDPNNHFDLSOPITKIALYKSGIIPVVRVRCMKRSGIRFTINGH 119
  |||||||
QY 137 STFNILVLTNVGAGDVHSVSKSGRTKQMLSRNMGQMSNSYLNGSLSFVTTSDR 196
  |||||||
Db 120 SYFNILVLTNVGAGDVHSVSKSGRTKQMLSRNMGQMSNSYLNGSLSFVTTSDR 179
  |||||||
QY 197 RSVSPFNAPPMPSFGOTTTGCGQFR 222
  |||||||
Db 180 RSVLSTNAPPMPSFGOTTTGCGQFR 205
  |||||||

```

```

RESULT 3
US-10-219-999-42752
: Sequence 42752, Application US/10219999
: GENERAL INFORMATION:
: APPLICANT: Cao, Yongwei
: APPLICANT: Edgerton, Michael D
: APPLICANT: Hinkle, Gregory J.
: APPLICANT: Kovalic, David K.
: APPLICANT: Liu, Jingdong
: APPLICANT: Stein, Joshua
: TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
: FILE REFERENCE: 38-10(52726)C
: CURRENT APPLICATION NUMBER: US/10/219,999
: PRIOR FILING DATE: 2002-08-15
: PRIOR APPLICATION NUMBER: US 60/324,109
: PRIOR FILING DATE: 2001-09-21
: PRIOR APPLICATION NUMBER: US 60/312,544
: PRIOR FILING DATE: 2001-08-15
: NUMBER OF SEQ ID NOS: 63520
: SEQ ID NO 42752
: LENGTH: 264
: TYPE: PRT
: ORGANISM: Glycine max
: US-10-219-999-42752

```

```

Query Match 75.9%; Score 913.5; DB 6; Length 264;
Best Local Similarity 73.6%; Pred. No. 3.1e-88;
Matches 167; Conservative 19; Mismatches 32; Indels 9; Gaps 2;
QY 3 GPWINAHATFYXXGDAXXTGACGAGYGNLYSGYGLTAAALSTALFDOGLSCGACXELMC 62

```

```

Db 39 GGWTNAHATFYXXGDAXXTGACGAGYGNLYSGYGTNTALSTALFNNGLSCGSCYEIRK 98
  |||||||
QY 63 VNDPOMCIRKRSIVTATNFCPP-----GCACDPNNHFDLSOPITKIALYKSGII 114
  |||||||
Db 99 ANDHRKCLPG-STIVTATNFCPPNNALPNNNGCMCNPLQHDLPRLTAQYKAGIV 157
  |||||||
QY 115 PVMTRVRCRSGGIRFTINGHSYFNILVLTNVGAGDVHSVSKSGRTKQMLSRNMGQ 174
  |||||||
Db 158 PVMTRVRCRSGGIRFTINGHSYFNILVLTNVGAGDVHSVSKSGRTKQMLSRNMGQ 217
  |||||||
QY 175 NMOSNSYLNGSLSFVTTSDRSVSNFVAPPMPSFGOTTTGCGQFR 221
  |||||||
Db 218 NMOSNSYLNGSLSFVTTSDRSVSNFVAPPMPSFGOTTTGCGQFR 264
  |||||||

```

```

RESULT 4
US-09-935-625-2250
: Sequence 2250, Application US/09935625
: GENERAL INFORMATION:
: APPLICANT: N. ALEXANDROV et al.
: TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPA
: FILE REFERENCE: 2750-1481P
: CURRENT APPLICATION NUMBER: US/09/935,625
: CURRENT FILING DATE: 2001-08-24
: NUMBER OF SEQ ID NOS: 33136
: SEQ ID NO 2250
: LENGTH: 236
: TYPE: PRT
: ORGANISM: Arabidopsis thaliana
: FEATURE:
: NAME/KEY: peptide
: LOCATION: 1..236
: OTHER INFORMATION: Ceres Seq. ID no. 3446952
: US-09-935-625-2250

```

```

Query Match 75.5%; Score 909.5; DB 5; Length 236;
Best Local Similarity 73.1%; Pred. No. 7.1e-88;
Matches 166; Conservative 19; Mismatches 33; Indels 9; Gaps 2;
QY 3 GPWINAHATFYXXGDAXXTGACGAGYGNLYSGYGLTAAALSTALFDOGLSCGACXELMC 62
  |||||||
Db 11 GGWTNAHATFYXXGDAXXTGACGAGYGNLYSGYGTNTALSTALFNNGLSCGACPEIRK 70
  |||||||
QY 63 VNDPOMCIRKRSIVTATNFCPP-----GCACDPNNHFDLSOPITKIALYKSGII 114
  |||||||
Db 71 QNDGKCLPG-STIVTATNFCPPNNALPNNAGCMCNPPQHFDSOPVFORIAQYRAGIV 129
  |||||||
QY 115 PVMTRVRCRSGGIRFTINGHSYFNILVLTNVGAGDVHSVSKSGRTKQMLSRNMGQ 174
  |||||||
Db 130 PVMTRVRCRSGGIRFTINGHSYFNILVLTNVGAGDVHSVSKSGRTKQMLSRNMGQ 189
  |||||||
QY 175 NMOSNSYLNGSLSFVTTSDRSVSNFVAPPMPSFGOTTTGCGQFR 221
  |||||||
Db 190 NMOSNSYLNGSLSFVTTSDRSVSNFVAPPMPSFGOTTTGCGQFR 236
  |||||||

```

```

RESULT 5
US-09-935-625-2254
: Sequence 2254, Application US/09935625
: GENERAL INFORMATION:
: APPLICANT: N. ALEXANDROV et al.
: TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPA
: FILE REFERENCE: 2750-1481P
: CURRENT APPLICATION NUMBER: US/09/935,625
: CURRENT FILING DATE: 2001-08-24
: NUMBER OF SEQ ID NOS: 33136
: SEQ ID NO 2254
: LENGTH: 236
: TYPE: PRT
: ORGANISM: Arabidopsis thaliana

```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 11, 2002, 14:57:45 ; Search time 31.293 Seconds
(Without alignments)
1988.637 Million cell updates/sec

Title: US-09-896-301-3
Perfect score: 1204
Sequence: 1 HMGWIMAHATFYXGDAAX.....NVAPTWFGQTYTGQFRY 222

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1020648 seqs, 280317128 residues

Total number of hits satisfying chosen parameters: 1020648

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Pending Patents_AA_New:**
1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep:**
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep:**
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep:**
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep:**
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:**
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep:**
7: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep:**

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	1004.5	83.4	247	US-10-219-999-35819	Sequence 35819, A
2	952.5	79.1	205	US-10-219-999-59261	Sequence 59261, A
3	913.5	75.9	264	US-10-219-999-42752	Sequence 42752, A
4	909.5	75.5	236	US-09-935-625-2250	Sequence 2250, Ap
5	909.5	75.5	236	US-09-935-625-2254	Sequence 2254, Ap
6	909.5	75.5	236	US-09-935-625-23150	Sequence 23150, A
7	909.5	75.5	236	US-09-935-625-23154	Sequence 23154, A
8	909.5	75.5	236	US-09-935-625-31516	Sequence 31516, A
9	909.5	75.5	236	US-09-935-625-31520	Sequence 31520, A
10	909.5	75.5	250	US-09-935-625-2249	Sequence 2249, Ap
11	909.5	75.5	250	US-09-935-625-2253	Sequence 2253, Ap
12	909.5	75.5	250	US-09-935-625-23149	Sequence 23149, A
13	909.5	75.5	250	US-09-935-625-23153	Sequence 23153, A
14	909.5	75.5	250	US-09-935-625-23155	Sequence 23155, A
15	909.5	75.5	250	US-09-935-625-31519	Sequence 31519, A
16	909.5	75.5	250	US-09-935-625-42838	Sequence 42838, A
17	909.5	75.5	261	US-09-935-625-2246	Sequence 2246, Ap
18	909.5	75.5	261	US-09-935-625-23146	Sequence 23146, A
19	893.5	74.2	261	US-09-935-625-31512	Sequence 31512, A
20	893.5	74.2	275	US-09-935-625-2245	Sequence 2245, Ap
21	893.5	74.2	275	US-09-935-625-23145	Sequence 23145, A
22	893.5	74.2	275	US-09-935-625-31511	Sequence 31511, A
23	868.5	72.1	277	US-10-219-999-36624	Sequence 36624, A
24	849.5	70.6	258	US-10-150-559-4	Sequence 4, Appl1
25	842	69.9	253	US-09-935-625-407	Sequence 407, App
26	842	69.9	253	US-09-935-625-11355	Sequence 11355, A

27	842	69.9	253	US-09-935-625-24404	Sequence 24404, A
28	838.5	69.6	273	US-10-219-999-36819	Sequence 36819, A
29	833.5	69.2	207	US-09-935-625-2251	Sequence 2251, Ap
30	833.5	69.2	207	US-09-935-625-2255	Sequence 2255, Ap
31	833.5	69.2	207	US-09-935-625-23151	Sequence 23151, A
32	833.5	69.2	207	US-09-935-625-23155	Sequence 23155, A
33	833.5	69.2	207	US-09-935-625-31517	Sequence 31517, A
34	833.5	69.2	207	US-09-935-625-31521	Sequence 31521, A
35	825.5	68.6	253	US-10-219-999-54448	Sequence 54448, A
36	820.5	68.1	258	US-10-219-999-42206	Sequence 42206, A
37	817.5	67.9	232	US-09-935-625-2247	Sequence 2247, Ap
38	817.5	67.9	232	US-09-935-625-23147	Sequence 23147, A
39	817.5	67.9	232	US-09-935-625-31513	Sequence 31513, A
40	817.5	67.9	262	US-10-219-999-37064	Sequence 37064, A
41	807.5	67.1	253	US-10-219-999-51167	Sequence 51167, A
42	805.5	66.9	253	US-60-370-796-9	Sequence 9, Appl1
43	790	65.6	207	US-09-935-625-408	Sequence 408, App
44	790	65.6	207	US-09-935-625-11356	Sequence 11356, A
45	790	65.6	207	US-09-935-625-24405	Sequence 24405, A

ALIGNMENTS

RESULT 1

US-10-219-999-35819

Sequence 35819, Application US/10219999

GENERAL INFORMATION:

APPLICANT: Cao, Yongwei

APPLICANT: Edgerton, Michael D

APPLICANT: Hinkle, Gregory J.

APPLICANT: Kovalic, David K.

APPLICANT: Liu, Jindong

APPLICANT: Stein, Joshua

TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT

FILE REFERENCE: 38-10(52726)C

CURRENT APPLICATION NUMBER: US/10/219, 999

PRIOR FILING DATE: 2002-08-15

PRIOR APPLICATION NUMBER: US 60/324,109

PRIOR FILING DATE: 2001-09-21

PRIOR APPLICATION NUMBER: US 60/312,544

PRIOR FILING DATE: 2001-08-15

NUMBER OF SEQ ID NOS: 63520

SEQ ID NO 35819

LENGTH: 247

TYPE: PRT

ORGANISM: Glycine max

US-10-219-999-35819

Query Match

Best Local Similarity 83.4%; Score 1004.5; DB 6; Length 247;

Matches 176; Conservative 22; Mismatches 21; Indels 1; Gaps 1;

3 GWINAHAFYXXGAXXMTMGACGYNLYSGYGHETALSTALFDGLSCGACXELMC U2

29 GGTNAHAFYFGSDSGMTGACGYNLYSGYGHETALSTALFDGLSCGACXELMC 88

63 VNDPQWICRISIVATNFCPPGACDPPNHHFDSQPIYKIALYSGIIPVYRRV 122

89 ANDPQWICRISIVATNFCPPGACDPPNHHFDSQPIYKIALYSGIIPVYRRV 147

123 CKRSGIRFTINGSHFNLVLTNNGGAGDVSHVSKGSRRTWOLMSRWGCONMSNYL 182

148 CMRRGIRFTINGSHFNLVLTNNGGAGDVSHVSKGSRRTWOLMSRWGCONMSNYL 207

183 NGOSLFVYTTSDRSVSNVFNAPPTWFGQTYTGQFRY 222

208 NGOSLFVYTTSDRSVSNVFNAPPTWFGQTYTGQFRY 247

RESULT 2

US-10-219-999-59261

Sequence 59261, Application US/10219999

Sun Oct 13 14:18:53 2002

us-09-896-301-3.rpr

Page 6

Search completed: October 11, 2002, 15:00:59
Job time : 11.141 secs

```
OY      3 GPMINAHATFYXXGDAAXXTMTMGACCGYNLYSOGYGLETALSTALEFDQSLCSGACKXELMC   62
        | | | | | | | | | | | | | | | | | | | | | : | | | | | | | | | | |
Db     34 GGGMOTLHNFVGGADATCTGTMGAGCCGYNLSOSGVTGSTALTALSTALENNLSISGAAYLEIRC   93
                                                | | | | | | | | | | | | | | |
OY      63 VNDPOMCIGRSIVTATTNPFR-----GGACDPPNHHEFLDSPITKIALTKSGII    114
        ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

A:Accession: F84831
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-253 <STO>
A:Cross-references: GB:AE002095; NID:g2651297; PIDDN:AA867577.1; GSPDH:GN001357
C:GeneLinks:
A:Gene: AE2940610
A:Map position: 2

QY 115 PVMRRVRCRSGGIRFTINGHSYFNLVLTNNGAGADVHVSMSKSRKWLMSRMNQ 174
 Db 127 PLYTRVPCLRKGGIRFTVNGHSYFNLVLTNNGAGADVHVSMSKSRKWLMSRMNQ 186
 QY 175 NMOSNSTLNGQSLSFVYTTSDRSYVSFNVAPPTWSFGQTYTGQ 219
 Db 187 NMOSNSTLNGQSLSFVYTTSDGRTYVSNNVAPSNMFGQTFEGSQ 231

RESULT 6

expansin (clone plexp2) - loblolly pine (fragment)

C:Species: Pinus taeda (loblolly pine)
 C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 26-May-2000
 C:Accession: T09818
 R:Hutchinson, K.W.; Singer, P.B.; Diaz-Sala, C.; Greenwood, M.S.
 submitted to the EMBL Data Library, July 1996
 A:Description: Expansins are conserved in conifers and expressed in response to exogenous A:Reference number: Z16866
 A:Accession: T09818
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-232 <HUT>
 A:Cross-references: EMBL:U64890; NID:q1778096; PID:q1778099
 A:Experimental source: clone plexp2; hypocotyl
 C:Superfamily: expansin

Query Match 71.7%; Score 863.5; DB 2; Length 232;
 Best Local Similarity 68.9%; Pred. No. 2.3e-73;
 Matches 155; Conservative 24; Mismatches 37; Indels 9; Gaps 2;

QY 3 GPMINAHATFYXXGDAXXTMGACGYNLSQGYGLETALSTALFDQGLSGCAXELMC 62
 Db 8 GGMESAHATFYGGSDASCTMGACGYNLSQGYGLETALSTALFDQGLSGCAXELMC 67
 QY 63 VNDPQWCKRSGIRFTVNGHSYFNLVLTNNGAGADVHVSMSKSRKWLMSRMNQ 114
 Db 68 NDDPQWCKLPG-TVTITATNCPNNALPNDNGGWCNPPLOHFDMAEPALAKYRGIV 126
 QY 115 PVMRRVRCRSGGIRFTINGHSYFNLVLTNNGAGADVHVSMSKSRKWLMSRMNQ 174
 Db 127 PLYTRVPCLRKGGIRFTVNGHSYFNLVLTNNGAGADVHVSMSKSRKWLMSRMNQ 186
 QY 175 NMOSNSTLNGQSLSFVYTTSDRSYVSFNVAPPTWSFGQTYTGQ 219
 Db 187 NMOSNSTLNGQSLSFVYTTSDGRTYVSNNVAPSNMFGQTFEGSQ 231

RESULT 7

expansin s1 precursor - cucumber

C:Species: Cucumis sativus (cucumber)
 C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
 C:Accession: T10079
 R:Shcherban, T.Y.; Shi, J.; Durachko, D.M.; Gullitman, M.J.; McQueen-Mason, S.J.; Shieh, Proc. Natl. Acad. Sci. U.S.A. 92, 9245-9249, 1995
 A:Title: Molecular cloning and sequence analysis of expansins--a highly conserved, multi A:Reference number: Z14894; MUID:96016146
 A:Accession: T10079
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-250 <SHC>
 A:Cross-references: EMBL:U30382; NID:q1040874; PID:AMB37746.1; PID:q1040875
 A:Experimental source: cultivar Burpee Pickler
 C:Genetics:
 A:Gene: EXP1
 C:Function:
 A:Description: mediates cell wall extension
 C:Superfamily: expansin
 F:1-33/Domain: signal sequence #status predicted <SIG>
 F:24-250/Product: expansin #status predicted <MAT>

Query Match 71.2%; Score 857.5; DB 2; Length 250;
 Best Local Similarity 67.3%; Pred. No. 9.2e-73;
 Matches 152; Conservative 26; Mismatches 39; Indels 9; Gaps 2;

QY 3 GPMINAHATFYXXGDAXXTMGACGYNLSQGYGLETALSTALFDQGLSGCAXELMC 62
 Db 26 GGMESAHATFYGGSDASCTMGACGYNLSQGYGLETALSTALFDQGLSGCAXELMC 85
 QY 63 VNDPQWCKRSGIRFTVNGHSYFNLVLTNNGAGADVHVSMSKSRKWLMSRMNQ 114
 Db 86 TNDPQWCKLPG-TVTITATNCPNNALPNDNGGWCNPPLOHFDMAEPALAKYRGIV 144
 QY 115 PVMRRVRCRSGGIRFTINGHSYFNLVLTNNGAGADVHVSMSKSRKWLMSRMNQ 174
 Db 145 PVSRRVPCMKKGVIRFTINGHSYFNLVLTNNGAGADVHVSMSKSRKWLMSRMNQ 204
 QY 175 NMOSNSTLNGQSLSFVYTTSDRSYVSFNVAPPTWSFGQTYTGQ 220
 Db 205 NMOSNSTLNGQSLSFVYTTSDGRTYVSNNVAPSNMFGQTFEGSQ 250

RESULT 8

expansin (clone plexp4) - loblolly pine (fragment)

C:Species: Pinus taeda (loblolly pine)
 C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 26-May-2000
 C:Accession: T09825
 R:Hutchinson, K.W.; Singer, P.B.; Diaz-Sala, C.; Greenwood, M.S.
 submitted to the EMBL Data Library, July 1996
 A:Description: Expansins are conserved in conifers and expressed in response to exogenous A:Reference number: Z16866
 A:Accession: T09825
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-232 <HUT>
 A:Cross-references: EMBL:U64892; NID:q1778102; PID:q1778103
 A:Experimental source: clone plexp4
 C:Superfamily: expansin

Query Match 71.1%; Score 856.5; DB 2; Length 232;
 Best Local Similarity 68.4%; Pred. No. 1e-72;
 Matches 154; Conservative 24; Mismatches 38; Indels 9; Gaps 2;

QY 3 GPMINAHATFYXXGDAXXTMGACGYNLSQGYGLETALSTALFDQGLSGCAXELMC 62
 Db 8 GGMESAHATFYGGSDASCTMGACGYNLSQGYGLETALSTALFDQGLSGCAXELMC 67
 QY 63 VNDPQWCKRSGIRFTVNGHSYFNLVLTNNGAGADVHVSMSKSRKWLMSRMNQ 114
 Db 68 NDDPQWCKLPG-TVTITATNCPNNALPNDNGGWCNPPLOHFDMAEPALAKYRGIV 126
 QY 115 PVMRRVRCRSGGIRFTINGHSYFNLVLTNNGAGADVHVSMSKSRKWLMSRMNQ 174
 Db 127 PLYTRVPCLRKGGIRFTVNGHSYFNLVLTNNGAGADVHVSMSKSRKWLMSRMNQ 186
 QY 175 NMOSNSTLNGQSLSFVYTTSDRSYVSFNVAPPTWSFGQTYTGQ 219
 Db 187 NMOSNSTLNGQSLSFVYTTSDGRTYVSNNVAPSNMFGQTFEGSQ 231

RESULT 9

probable expansin F13M22.14 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 16-Feb-2001
 C:Accession: T02530; C64795
 R:Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Syke submitted to the EMBL Data Library, June 1998
 A:Description: Arabidopsis thaliana chromosome II BAC F13M22 genomic sequence.
 A:Reference number: Z14677
 A:Accession: T02530
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 11, 2002, 14:53:05 ; Search time 11.141 Seconds

(without alignments)
1914.721 Million cell updates/sec

Title: US-09-896-301-3

Perfect score: 1204

Sequence: 1 HMGPMINAHATFYXXGDAXX.....NVAPPTWSEFGQYTTGGQPRY 222

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_71:*

SUMMARIES

Result No. Score Query Match Length DB ID Description

1	1195	99.3	255	2	T50655	expansin EXP5 [imp
2	909.5	75.5	237	2	T50654	expansin EXP1 [imp
3	907.5	75.4	248	2	C84444	probable expansin
4	870.5	72.3	232	2	T09821	expansin (clone pp
5	864.5	71.8	232	2	T09826	expansin (clone pp
6	863.5	71.7	232	2	T09818	expansin (clone pp
7	857.5	71.2	230	2	T10079	expansin SI precu
8	856.5	71.1	232	2	T09825	expansin (clone pp
9	855.5	71.1	262	2	T02530	probable expansin
10	852.5	70.8	258	2	T09786	expansin - upland
11	845.5	70.2	258	2	T48247	expansin - like prot
12	842	69.9	253	2	T84831	expansin-like prot
13	832.5	69.1	257	2	D84820	probable expansin
14	832.5	69.1	258	2	S53082	pollen allergen ho
15	825.5	68.6	260	2	T47689	expansin-like prot
16	823	68.4	255	2	T50656	expansin EXP2 [imp
17	816.5	67.8	257	2	T50658	expansin 9 [import
18	815.5	67.7	246	2	T04175	expansin - rice
19	805.5	66.9	251	2	T03258	expansin 2 - rice
20	804.5	66.8	259	2	T02727	probable expansin
21	804.5	66.8	259	2	T50653	expansin EXP6 [imp
22	799.5	66.4	262	2	T50660	alpha-expansin 2 l
23	792.5	65.8	255	2	T06573	expansin 18 - toma
24	788.5	65.5	260	2	T08016	probable expansin
25	782.5	65.0	261	2	T07630	expansin 1 - tomat
26	778.5	64.7	258	2	T10083	expansin S2 precu
27	776.5	64.5	264	2	T50659	alpha-expansin OSE
28	734.5	61.0	252	2	T86335	hypothetical prote
29	727.5	60.4	255	2	T02010	expansin homolog T

30	725	60.2	261	2	T03737	expansin - rice
31	698	58.0	255	2	T03299	expansin 3 - rice
32	624.5	51.9	257	2	G96654	hypothetical prote
33	620.5	51.5	257	2	F86259	protein T12C24.10
34	592.5	49.2	160	2	T09871	expansin - upland
35	548.5	45.6	256	2	T05648	expansin homolog F
36	278.5	23.1	259	2	T50657	expansin (clone pp
37	278.5	23.1	271	2	H84592	beta-expansin [imp
38	257.5	21.4	276	2	T09041	beta-expansin [imp
39	256	21.3	102	2	T09828	clml protein homol
40	252.5	21.0	277	2	S48032	expansin (clone pp
41	250	20.8	277	2	T09815	clml protein - soy
42	241	20.0	491	2	F96681	expansin (clone pp
43	237	19.7	261	2	T04301	protein F1E22.6 [i
44	225.5	18.7	263	2	S13614	beta-expansin - ri
45	224	18.6	259	2	E84886	major allergen lol
						probable beta-expa

ALIGNMENTS

RESULT 1
T50655
expansin EXP5 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 28-Jul-2000
C:Accession: T50655
R:Shcherban, T.Y.; Shi, J.; Durachko, D.M.; Gultinan, M.J.; McQueen-Mason, S.J.; Shi
Proc. Natl. Acad. Sci. U.S.A. 92, 9245-9249, 1995
A:Title: Molecular cloning and sequence analysis of expansins--a highly conserved, mu
A:Reference number: 214894; MUID:96016146
A:Accession: T50655
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-255 <SHC>
A:Cross-references: EMBL:U30478; PTDN:AAB38071.1
A:Experimental source: variety Columbia
A:Gene: Exps
C:Function:
A:Description: induces extension (creep) in plant cell walls
C:Superfamily: expansin
C:Keywords: cell wall

Query Match 99.3%; Score 1195; DB 2; Length 255;
Best Local Similarity 97.7%; Pred. No. 2; Se-104;
Matches 217; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 HMGPMINAHATFYXXGDAXXTMGACGYNLSQGYGLTAALSTALFDGLSGACXEL 60
Db 34 HMGPMINAHATFYGGGDSGTMGACGYNLSQGYGLTAALSTALFDGLSGACFEL 93

QY 61 MCVNDPWCICIGRSIVVTATNECPGACDPNNHFDLSQPIYEKIALYKSGIIPVYRR 120
Db 94 MCVNDPWCICIGRSIVVTATNECPGACDPNNHFDLSQPIYEKIALYKSGIIPVYRR 153

QY 121 VRCRSGGIPTINGHYFNLVLTNNGAGDVHSVSMKSRKRWQMSRNQMNQNS 160
Db 154 VRCRSGGIPTINGHYFNLVLTNNGAGDVHSVSMKSRKRWQMSRNQMNQNS 213

QY 181 YLNGOSLFFVTTSDRRSVSFNVAPTWSEFGQYTTGGQPRY 222
Db 214 YLNGOSLFFVTTSDRRSVSFNVAPTWSEFGQYTTGGQPRY 255

RESULT 2
T50654
expansin EXP1 [imported] - Arabidopsis thaliana (fragrant)
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 02-Sep-2000
C:Accession: T50654
R:Shcherban, T.Y.; Shi, J.; Durachko, D.M.; Gultinan, M.J.; McQueen-Mason, S.J.; Shi
Proc. Natl. Acad. Sci. U.S.A. 92, 9245-9249, 1995

DR PDB: 2PBH: 25-FEB-98.
 DR PDB: 3PBH: 25-FEB-98.
 DR MEROPS: C01.060: -
 DR SWISS-2DPAGE: P07858; HUMAN.
 DR MIM: 116810: -
 DR InterPro: IPR000668; Peptidase_C1.
 DR InterPro: IPR000169; Thiolprot_act_site.
 DR Pfam: PF00112; Peptidase_C1; 1.
 DR PRINTS: PR00705; PAPA1N.
 DR PROSITE: PS00139; THIOL_PROTEASE_CYS; 1.
 DR PROSITE: PS00639; THIOL_PROTEASE_HIS; 1.
 DR PROSITE: PS00640; THIOL_PROTEASE_ASN; 1.
 DR Hydrolase: Thiol protease; lysosome; glycoprotein; zymogen; signal;
 KW Polymorphism: 3D-structure.
 KM SIGNAL: 1 17
 FT PROPEP 18 79
 FT CHAIN 80 79
 FT CHAIN 80 126
 FT CHAIN 129 333
 FT PROPEP 334 339
 FT ACT_SITE 108 108
 FT ACT_SITE 278 278
 FT ACT_SITE 298 298
 FT ACT_SITE 93 122
 FT DISULFID 105 150
 FT DISULFID 141 207
 FT DISULFID 142 146
 FT DISULFID 179 211
 FT DISULFID 187 198
 FT CARBOHYD 192 192
 FT VARIANT 26 26
 FT CONFLICT 228 228
 FT SEQUENCE 339 AA: 37807 MW: 0605FE184AE47070 CRC64:
 N-LINKED (GLCNAC. . .).
 V -> L.
 /FTID=VAR_006724.
 N -> D (IN REF 4).
 Query Match 7.0%; Score 84; DB 1; Length 339;
 Best Local Similarity 19.8%; Pred. No. 0.9;
 Matches 44; Conservative 24; Mismatches 80; Indels 74; Gaps 10;

QY 7 NAHATF-YXXGDAXXTWGACGYGNLYSGYGLT-----AALSTALFDGSLGACGX 58
 DB 126 NAHVSVEVSAEDLLTCCGSGMGDG--CNGCYPAEAMNFTKRLGLVSGGLYSHVGCRRPS 183
 QY 59 ELWC---VNDPOMCIKRSIVTATNFCRPG-----GACDRPHHFDLSQPT 102
 DB 184 IPCEHHVNGSRPCTEGTPEKSKICERGVSPYTKQDKHGYNSYSNSEKDIMAEI 243
 QY 103 YEK-----IALYKSGIIPVYRVRCKRSGGIRFTINGHSYENLVLTNVGCA 150
 DB 244 YKNGPVEGAFSVSDPLLYKSGV---YQHVTGEMMG----- 277
 QY 151 GDVHSVSMKG---SRTKQMLSMRNGMGNONSYS---LNGQ 185
 DB 278 ---HAIRILGWGVENGTPYVLVANSWMTDMGDNCFKILKQ 316
 RESULT 12
 ID HRA1_MOUSE STANDARD: PRT: 480 AA.
 AC Q9RL18:
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Serine protease HTRA1 precursor (EC 3.4.21.-).
 CN PRS11 OR HTRA1 OR HTRA.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=ICR; TISSUE=Brain;
 RA 'Oka C., Some A., Kanda H., Kawachi M.;

RT "The role of murine serine protease HTRA in osteogenesis.";
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Protease that regulate the availability of IGRs by
 CC cleaving IGF-binding proteins (by similarity).
 CC -1- SUBCELLULAR LOCATION: Secreted (by similarity).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S2C; ALSO KNOWN AS THE
 CC DSGP/DEGP/DEGS FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 IGFBP DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 KAZAL-LIKE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 PDZ/DHR DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF172994; AAD49422.1; -
 CC MGD: MGI:1929076; Pss11.
 DR InterPro: IPR000867; IGFBP.
 DR InterPro: IPR001478; PDZ.
 DR InterPro: IPR001940; Protease2C.
 DR InterPro: IPR001254; Trypsin.
 DR InterPro: IPR002350; kazal.
 DR Pfam: PF00219; IGFBP; 1.
 DR Pfam: PF00050; kazal; 1.
 DR Pfam: PF00595; PDZ; 1.
 DR Pfam: PF00089; trypsin; 1.
 DR Pfam: PRO00834; PROTEASES2C.
 DR SMART: SM00121; IB; 1.
 DR SMART: SM00280; KAZAL; 1.
 DR SMART: SM00288; PDZ; 1.
 DR PROSITE: PS50106; PDZ; 1.
 DR KW Hydrolase; Serine protease; Growth factor binding; signal.
 FT SIGNAL 1 22
 FT CHAIN 23 480
 FT DOMAIN 37 94
 FT DOMAIN 101 155
 FT DOMAIN 204 364
 FT DOMAIN 365 467
 FT ACT_SITE 220 220
 FT ACT_SITE 250 250
 FT ACT_SITE 328 328
 FT SEQUENCE 480 AA: 51246 MW: 54BB9BA6C99A7BF4 CRC64:
 Query Match 6.7%; Score 81; DB 1; Length 480;
 Best Local Similarity 21.9%; Pred. No. 2.6;
 Matches 48; Conservative 26; Mismatches 73; Indels 72; Gaps 11;
 QY 16 GDAXXTMGACGYGNLYSGYGLTALSTALFDGSLGACXELMCVNDPOMCIKRSI 75
 DB 73 GAACGLOGPCGEGLOCVLPGVASATVRRRAGACVCASSEPVCGSDAK----- 124
 QY 76 VVTATNFC-----PP-----GGAC-----DPN---HHFDLSQPTYEKIA-- 107
 DB 125 --TYTNLCOLRAASRRSEKILPPQPVIVLORACGCGEDPNSLRHKYNTADVVEFAFD 182
 QY 108 -----LYSGIIPVYRVRCKRSGGIRFTINGHSYENLVLTNVGCGAGDVHSVSMKGS 162
 DB 183 VVKHELKRR--LPSKREVVPVASSGGLVSDG-----LIVTN-----AHVVTNK-NR 227
 QY 163 TKQMLSMRNGMGNONSYSYNGQSLSPVYTTSDRRSVS 201
 DB 228 VKVELK-----NCATYEALIKVDKADIA 252
 RESULT 13
 ID CYPH_BIAGE STANDARD: PRT: 164 AA.
 AC P54985;
 DT 01-OCT-1996 (Rel. 34, Created)

or send an email to license@slsb.ch.

CC EMBL: AI133035; CAB37075.1; -
 CC EMBL: AE001585; AAD18163.1; -
 DR EMBL: AE002237; AAF38570.1; -
 DR EMBL: AE002545; BAA98215.1; -
 DR TIGR: CP0770; -
 DR InterPro: IPR003368; DUF145.
 DR InterPro: IPR003357; OMP.
 DR Pfam: PF02415; DUF145; 1.
 DR Pfam: PF02385; OMP; 1.
 KW Outer membrane; Signal; Multigene family; Complete proteome.
 FT SIGNAL 1 26 POTENTIAL.
 FT CHAIN 27 922 PROBABLE OUTER MEMBRANE PROTEIN PMPL.
 FT CONFLICT 14 14 F -> L (IN REF. 1).
 FT CONFLICT 375 375 Y -> C (IN REF. 1).
 FT CONFLICT 606 606 D -> N (IN REF. 1).
 FT SEQUENCE 922 AA; 100457 MW; DFF2AB6333AB031C CRC64;
 SQ
 Query Match 7.0%; Score 84.5; DB 1; Length 922;
 Best Local Similarity 24.2%; Pred. No. 2.4;
 Matches 59; Conservative 29; Mismatches 81; Indels 75; Gaps 15;

QY 6 INAAHFFXXGDAXXTMGAGCYGLYSQG-----YGL--ETAAALSTALPDQGLSCG 55
 DB 206 VNAQETFRQNTAKNGSSGA-----LYSDGDIDIDONAYVLFRENEALLTAIGKGAVC- 259
 QY 56 ACXELMCVNDPQWCIKRSIVVTATNFCPCGACDP--PNHHF-DLSQPIEKALYKSGT 113
 DB 260 -----CL-----PTSGSTPVPYTFESDNKOLVERNHSTMG- 292
 QY 114 IPVWYRVRCKRRSGGIRTINGHYENLVLTNVGA-----GDVH-----SVSMKGS 161
 DB 293 -GAIYARLSTISGSGPTLFNNISYAN---SONIGCAIADTGGELISLSEKGTITFGQ 348
 QY 162 RTKQMLSRNMGWQMSNSYL-----NGOSLSF---VVTSDKRSVSEFVAPPTTSFQ 213
 DB 349 RTSLPFL--NGIHLLQNAKFLKIDARNGYSIEFYDPTTSEADSGSTQNLINGDK---NK 402
 QY 214 TYTG 217
 DB 403 EYTG 406

RESULT 11
 CATE_HUMAN
 ID CATE_HUMAN STANDARD; PRT; 339 AA.
 AC P07858;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cathepsin B precursor (EC 3.4.22.1) (Cathepsin B1) (APP secretase).
 GN CTSB OR CTSB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87017021; PubMed=3463996;
 RA Chan S.-J., San Segundo B., McCormick M.B., Steiner D.F.;
 RT "Nucleotide and predicted amino acid sequences of cloned human and
 mouse preprocathepsin B cDNAs.";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:7721-7725(1986).
 RP [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=gastric carcinoma;
 RX MEDLINE=94156194; PubMed=8112600;
 RA Cao L., Taggart R.T., Berguin I.M., Moln K., Fong D., Sloane B.F.;
 RT "Human gastric adenocarcinoma cathepsin B: isolation and sequencing
 of full-length cDNAs and polymorphisms of the gene.";
 RL Gene 139:163-169(1994).
 RN [3]

RP SEQUENCE OF 80-126 AND 129-333.
 RC TISSUE=Liver;
 RX MEDLINE=85127484; PubMed=3972105;
 RA Ritonja A., Popovic T., Turk V., Wiedenmann K., Machleidt W.;
 RT "Amino acid sequence of human liver cathepsin B.";
 RL FEBS Lett. 181:169-172(1985).
 RN [4]
 RP SEQUENCE OF 131-339 FROM N.A.
 RX MEDLINE=86206063; PubMed=3010323;
 RA Fong D., Calhoun D.H., Hsieh W.-T., Lee B., Wells R.D.;
 RT "Isolation of a cDNA clone for the human lysosomal proteinase
 cathepsin B.";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:2909-2913(1986).
 RP [5]
 RP SEQUENCE OF 80-91 AND 129-139.
 RC TISSUE=Liver;
 RX MEDLINE=92344620; PubMed=1637335;
 RA Moln K., Day N.A., Sameni M., Hasnain S., Hiram T., Sloane B.F.;
 RT "Human tumour cathepsin B. Comparison with normal liver cathepsin B.";
 RL Biochem. J. 285:427-434(1992).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (2.15 ANGSTROMS).
 RX MEDLINE=91330854; PubMed=1868826;
 RA Husli D., Zucic D., Turk D., Eng R.A., Mayr I., Huber R., Popovic T.,
 RT Turk V., Towatari T., Katunuma N., Bode W.;
 RT "The refined 2.15 A x-ray crystal structure of human liver cathepsin
 B: the structural basis for its specificity.";
 RL EMBO J. 10:2321-2330(1991).
 RN [7]
 RP X-RAY CRYSTALLOGRAPHY (3.2 ANGSTROMS).
 RX MEDLINE=96197789; PubMed=8617355;
 RA Turk D., Podobnik M., Kuhnelt R., Dollner M., Turk V.;
 RT "Crystal structures of human procathepsin B at 3.2- and 3.3-A
 resolution reveal an interaction motif between a papain-like cysteine
 protease and its propeptide.";
 RL FEBS Lett. 384:211-214(1996).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
 RX MEDLINE=97446326; PubMed=9299326;
 RA Podobnik M., Kuhnelt R., Turk V., Turk D.;
 RT "Crystal structure of the wild-type human procathepsin B at 2.5-A
 resolution reveals the native active site of a papain-like cysteine
 protease zymogen.";
 RL J. Mol. Biol. 271:774-788(1997).
 RN [9]
 RP -I- FUNCTION: THIOL PROTEASE WHICH IS BELIEVED TO PARTICIPATE IN
 INTRACELLULAR DEGRADATION AND TURNOVER OF PROTEINS. HAS ALSO
 BEEN IMPLICATED IN TUMOR INVASION AND METASTASIS.
 CC -I- CATALYTIC ACTIVITY: HYDROLASES PROTEINS, WITH A SPECIFICITY
 RESEMBLING THAT OF PAPAIN.
 CC -I- SUBUNIT: DIMER OF A HEAVY CHAIN AND A LIGHT CHAIN CROSS-LINKED
 BY A DISULFIDE BOND.
 CC -I- SUBCELLULAR LOCATION: Lysosomal.
 CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1; ALSO KNOWN AS THE
 PAPAIN FAMILY OF THIOL PROTEASES.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@slsb.ch).

DR EMBL: M14221; AAA52129.1; -
 DR EMBL: L16510; AAC37547.1; -
 DR EMBL: M13230; AAA52125.1; -
 DR PIR: A26496; KHHUB.
 DR PIR: A25432; A25432.
 DR PIR: S16513; S16513.
 DR PIR: S16514; S16514.
 DR PDB: 1HUC; 26-JAN-95.
 DR PDB: 1CSB; 03-APR-96.
 DR PDB: 1PBH; 23-FEB-98.

Db	75	GKCYLVTSTGGA	PCSSCGTGGAA	GSILVWYTNL	CCPNNGNM	QWCVVGT	NTQYGSY	HFD	134
Oy	98	L	98						
Db	135	I	135						
RESULT 9									
CART_MOUSE									
ID	CART_MOUSE	STANDARD;	PRT;	333	AA.				
AC	Q9R014	Q9WV51;							
DT	16-OCT-2001	(Rel. 40, Created)							
DT	16-OCT-2001	(Rel. 40, Last sequence update)							
DT	16-OCT-2001	(Rel. 40, Last annotation update)							
DE	Cathepsin J precursor	(EC 3.4.22.-) (Cathepsin P) (Catlp-p).							
GN	CTSJ	OR CTSP.							
OS	Mus musculus	(Mouse).							
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
CC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.								
OX	NCBI_TaxId=10090;								
RP	(1)								
RP	SEQUENCE FROM N.A.								
RC	STRAIN=C57BL/6J; TISSUE=Embryo;								
RX	MEDLINE=99456833; PubMed=10526153;								
RA	Tsai J et al., Deusing J., Peters C.;								
RT	"Cathepsin J, a novel murine cysteine protease of the papain family								
RL	with a placenta-restricted expression."								
RL	FEBS Lett. 459:299-304(1999).								
RN	(2)								
RP	SEQUENCE FROM N.A.								
RC	STRAIN=C57BL/6; TISSUE=Placenta;								
RA	Soi-Church K., French J., Troeber D., Mason R.W.;								
RT	"Cloning of a mouse cysteine protease."								
RL	Submitted (JUN-1999) to the EMBL/Genbank/DBJ databases.								
CC	-1- SUBCELLULAR LOCATION: Lysosomal (Potential).								
CC	-1- TISSUE SPECIFICITY: EXPRESSED IN PLENTA.								
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1; ALSO KNOWN AS THE								
CC	PAPAIN FAMILY OF THIOL PROTEASES.								
CC	-----								
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration								
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -								
CC	the European Bioinformatics Institute. There are no restrictions on its								
CC	use by non-profit institutions as long as its content is in no way								
CC	modified and this statement is not removed. Usage by and for commercial								
CC	entities requires a license agreement (See http://www.isb.ch/announce/								
CC	or send an email to license@isb-sib.ch).								
CC	-----								
DR	EMBL: AF136272; AAF13142.1; -								
DR	EMBL: AF158182; AADA1898.1; -								
DR	HSSP: P07711; ICDL.								
DR	MEROPS: C01.038; -								
DR	MCD: MGI:1349426; Ctsj.								
DR	InterPro: IPR000668; Peptidase_C1.								
DR	InterPro: IPR000169; Thiolprot_act_site.								
DR	Pfam: PF00112; Peptidase_C1; 1.								
DR	PRINTS: PR00705; PAPA1N.								
DR	PROSITE: PS00139; THIOL_PROTEASE_CYS; 1.								
DR	PROSITE: PS00639; THIOL_PROTEASE_HIS; 1.								
DR	PROSITE: PS00640; THIOL_PROTEASE ASN; 1.								
KW	Hydrolase; Thiol protease; Glycoprotein; Lysosome; Zymogen; Signal.								
FT	SIGNAL	1	17						
FT	PROPEP	18	112						
FT	CHAIN	113	333						
FT	ACT_SITE	137	137						
FT	ACT_SITE	275	275						
FT	ACT_SITE	299	299						
FT	DISULFID	134	177						
FT	DISULFID	168	210						
FT	DISULFID	268	321						
FT	CARBOHYD	71	71						
FT	CARBOHYD	216	216						
FT	CARBOHYD	220	220						

[illegible]

OS Phalaris aquatica (Canary grass).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 CC Poaceae; Phalaris.
 CC NCBI_TaxID=28479;
 [1]
 RN SEQUENCE FROM N.A.
 RC TISSUE-Pollen;
 RX MEDLINE=96105569; Pubmed=8564724;
 RA Suphloglu C., Singh M.B.;
 RT "Cloning, sequencing and expression in Escherichia coli of Pha a 1
 RT and four isoforms of Pha a 5, the major allergens of canary grass
 RT pollen."; Allergy 25:853-865(1995).
 RL Clin. Exp. Allergy 25:853-865(1995).
 RN [2]
 RP SEQUENCE OF 30-49.
 RX MEDLINE=93319091; Pubmed=7687099;
 RA Suphloglu C., Singh M.B., Simpson R.J., Ward L., Knox R.B.;
 RT "Identification of canary grass (Phalaris aquatica) pollen allergens
 RT by immunoblotting: IgE and IgG antibody-binding studies.";
 RL Allergy 48:273-281(1993).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- DISEASE: CAUSES GRASS POLLEN ALLERGY.
 CC -1- SIMILARITY: BELONGS TO THE LOL P I FAMILY OF ALLERGENS.
 CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE EG45 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE CBD DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: S80654; AAB35984.1; -
 DR HSSP: P43214; IMHO.
 DR InterPro: IPR000882; Pollen_allergen.
 DR Pfam: PF01357; Pollen_allergen; 1.
 DR PROSITE: PD002179; Pollen_allergen; 1.
 DR PROSITE: PS50843; EXPANSIN_CBD; 1.
 DR PROSITE: PS50842; EXPANSIN_EG45; 1.
 KW Allergen; Glycoprotein; Signal.
 FT SIGNAL 1 29
 FT CHAIN 30 269 MAJOR POLLEN ALLERGEN PHA A 1.
 FT DOMAIN 67 173 EXPANSIN-LIKE EG45.
 FT DOMAIN 187 268 EXPANSIN-LIKE CBD.
 FT CARBOHYD 38 38 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CONFLICT 37 37 P -> G (IN REF. 2).
 SQ SEQUENCE 269 AA; 29011 MW; 05835A58ACE877F CRC64;
 Query Match 16.8%; Score 202.5; DB 1; Length 269;
 Best Local Similarity 28.5%; Pred. No. 3,4e-12;
 Matches 65; Conservative 36; Mismatches 90; Indels 37; Gaps 11;
 Oy 5 WINAHAFYX--KDXAXXTMGACGCGYGLYGLTALSTALFDOGLSCGACXELMC 62
 Db 47 WLDKSTWYGKPTGAGPKDNGAGCYKDVAKPFGNGMTGCGNTPIFKDGCGSCFELKC 106
 Oy 63 VNDPWCICKGRSIVVATN-----FCPPGGACDPRNHHPSQIYKIA-----LYK 110
 Db 107 -SKPESC-SGEPITVHTDDNEPIAP-----YHFDLSGHAFSGSMKKGEENVRG 155
 Oy 111 SGIIPIVYRRVRCRSGGIRFTIN---GHSYFNLVLTNVGAGDVHYSMK-GSRTKW 165
 Db 156 AGELELGFRRVCKYKPGDTKPTFHVEKGSNPNYLALLVTKYVDGDDGVAAVDIKKGDKW 215
 Oy 166 QLMRRNMGNQNS--YLNQGSLSFYV--TTSDRKSVFNNVAPPW 209
 Db 216 IELKESWGAIMRIDTPKLTG--PFTVRYTTEGCTKAEEFEDVIEGCM 260

RESULT 6
 ID MPOL_ORYSA STANDARD; PRT; 263 AA.
 AC 040638;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Major pollen allergen Ory s 1 precursor (Ory s 1).
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Eriartoideae; Oryzae; Oryza.
 CC NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Anther;
 RX MEDLINE=96069591; Pubmed=7590339;
 RA Xu H., Theerakulpisut P., Goulding N., Suphloglu C., Singh M.B.,
 RA Bhalla P.L.;
 RT "Cloning, expression and immunological characterization of Ory s 1,
 RT the major allergen of rice pollen.";
 RL Gene 164:255-259(1995).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN MATURE ANTHERS BUT NOT IN
 CC VEGETATIVE OR OTHER FLORAL TISSUES.
 CC -1- DISEASE: CAUSES GRASS POLLEN ALLERGY.
 CC -1- SIMILARITY: BELONGS TO THE LOL P I FAMILY OF ALLERGENS.
 CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE EG45 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE CBD DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U31771; AAB6533.1; -
 DR HSSP: P43214; IMHO.
 DR InterPro: IPR000882; Pollen_allergen.
 DR Pfam: PF01357; Pollen_allergen; 1.
 DR PROSITE: PD002179; Pollen_allergen; 1.
 DR PROSITE: PS50843; EXPANSIN_CBD; 1.
 DR PROSITE: PS50842; EXPANSIN_EG45; 1.
 KW Allergen; Glycoprotein; Signal; Multigene family.
 FT SIGNAL 1 23
 FT CHAIN 24 263 MAJOR POLLEN ALLERGEN ORY S 1.
 FT DOMAIN 61 164 EXPANSIN-LIKE EG45.
 FT DOMAIN 178 259 EXPANSIN-LIKE CBD.
 FT CARBOHYD 32 32 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 263 AA; 28497 MW; B1C5F24EA39BD60 CRC64;
 Query Match 15.7%; Score 188.5; DB 1; Length 263;
 Best Local Similarity 31.3%; Pred. No. 7.2e-11;
 Matches 63; Conservative 29; Mismatches 82; Indels 27; Gaps 11;
 Oy 23 GGACGCGNYISGY-GLETAALSTALFDOGLSCGACXELMCVNDPWC-KGRSIVVAT 80
 Db 61 GGACGKVDVAKPFLGAMNSCG-NDPFLFKDGCKGCGSFEIKC-SKPEACSDKALIHVDM 118
 Oy 81 NCPGPGACDPR--NHFDLSQPIYKIA---LYKSGIIPVYRRVRCRSGGIRFTIN 134
 Db 119 N-----DEPIAVNHFDLSGLAMAKDKDELRKAGIIDPQFRVCKYPADRKIFPH 170
 Oy 135 ----GHSYFNLVLTNVGAGDVHYSM--KSRTRKQMLSMNMGNQNS--YLNQGS 186
 Db 171 IEKASPNYLLALLVTKYVADGDVVEVEIKEGSE-EMKALKESWGAIMRIDTPKPKLGP 229
 Oy 187 LSFVYTTSDRRSVSVFNVAP 207
 Db 230 SVRVTTGGAARRSSAEDAIIDP 250

RESULT 2	MPCL_CYNDA	MPCL_CYNDA	STANDARD;	PRT;	246 AA.
AC	MPCL_CYNDA				
AC	004701;				
DT	16-OCT-2001	(Rel. 40, Created)			
DT	16-OCT-2001	(Rel. 40, Last sequence update)			
DT	01-MAR-2002	(Rel. 41, Last annotation update)			
DE	Major pollen allergen Cyn d 1.				
GN	CYN01.				
OS	Cynodon dactylon (Bermuda grass).				
OC	Eukaryota: Viridiplantae: streps.				
OC	Spermatophyta: Magnoliophyta: Embryophyta: Tracheophyta:				
OC	Chloridoideae: Magnoliophyta: Liliopsida: Poales: Poaceae: PACC clade				
ON	NCBI_TaxID=28909;				
ON	CYN01.				

ID	MPC1_CYNDA	STANDARD;	PRT;	246 AA
AS	004701			

NC 004701;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Major pollen allergen Cyn d 1.
 GN CYN D1.
 OS Cynodon dactylon (Bermuda grass).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
 OC Chloridaceae; Cynodonteae; Cynodon.
 OX NCBI_TaxID=28909;
 [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TTSSUB-Pollen:
 RX MEDLINE=96347957; PubMed=8757211;
 RA Smith P.M., Suppholud C., Griffith I.J., Theriault K., Knox R.B.,
 RA Singh M.B.;
 RT Cloning and expression in yeast *Pichia pastoris* of a biologically
 RT active form of Cyn d 1, the major allergen of Bermuda grass pollen.;
 RT J. Allergy Clin. Immunol. 98:331-343(1996).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- DISEASE: CAUSES GRASS POLLEN ALLERGY.
 CC -1- SIMILARITY: BELONGS TO THE LOX P I FAMILY OF ALLERGENS.
 CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE EC45 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE CBD DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.elsdb.ch/announce/>
 CC or send an email to license@elsdb-sib.ch).
 CC -----

DR	InterPro:	IPR000883;	Pollen_allergen.	
DR	Pfam:	PF01357;	Pollen_allergen; 1.	
DR	PRINTS:	PRO1225;	EXPANSIN-FAMILY.	
DR	ProDom:	PD002179;	Pollen_allergen; 1.	
DR	ProSITE:	PS00843;	EXPANSIN_CBD; 1.	
DR	ProSITE:	PS00842;	EXPANSIN_EC45; 1.	
DR	ALLERGEN.			
KT	DOMAIN	39	145	EXPANSIN-LIKE EC45.
FT	DOMAIN	155	240	EXPANSIN-LIKE CBD.
ET	CARBOHYD	9	9	N-LINKED (GLCNAC....) (POTENTIAL)
QO	SEQUENCE	246 AA:	2688 MW:	43BD442DBA588322 CRC64;

Query Match	17.9%	Score 216;	DB 1;	Length 246;
Best local Similarity	28.1%;	Pred. No. 1.6e-13;		
Matches 66;	Conservative 37;	Mismatches 96;	Indels 36;	Gaps 10

RESULT 3			
MP_L1_LOLPR			
ID	MP_L1_LOLPR	STANDARD:	PRT: 263 AA.
AC	P14946;		
DT	01-APR-1990 (Rel. 14, Created)		
DT	01-AUG-1991 (Rel. 19, Last sequence update)		
DT	01-MAR-2002 (Rel. 41, Last annotation update)		
DE	Polien allergen Lol p 1 precursor (Lol p I) (allergen R7).		
OS	Lolium perenne (Perennial ryegrass).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;		
OC	Poaceae; Lolium.		
OX	NCBI_TaxID=4522;		

ID	MPPL_LOLPR	STANDARD;	PRT;	263 AA
1	11000			

AC P14946; P19964;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-AUG-1991 (Rel. 19, Last annotation update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Pollen allergen Lol p 1 precursor (Lol p 1) (Allergen R7).
OS Lolium perenne (Perennial ryegrass).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
OC Poaceae; Lolium.
OX NCBI_TaxID=4522;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90375479; PubMed=1697854;
RA Perez M., Ishioka G.Y., Walker L.E., Chesnut R.W.;
RT "CDNA cloning and immunological characterization of the rye grass
RT allergen Lol p I.";
RL J. Biol. Chem. 265:16210-16215(1990).
RN [2]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 24-48.
RX TISSUE=Pollen;
RC MEDLINE=91160716; PubMed=2001733;
RA Griffith I.J., Smith P.M., Pollock J., Theerakulpisut P.,
RA Avjoglou A., Davies S., Hough T., Singh M.B., Simpson R.J., Ward L.D.,
RA Knox R.B.;
RT "Cloning and sequencing of Lol pI, the major allergenic protein of
RT rye-grass pollen.";
RL FEBS Lett. 279:210-215(1991).
RN [3]

RC TISSUE=pollen;

RA MEDLINE=86242068; PubMed=3718465;
RX Cottam G.P., Moran D.M., Stauding R.;
RA Physicochemical and immunochemical characterization of allergenic
RT proteins from ryegrass (Lolium perenne) pollen prepared by a rapid
RT Biochem. J. 234:305-310(1986).
RN [4]
RP SEQUENCE OF 336-263.
RX MEDLINE=89364850; PubMed=2475768;
RA Esch R.E., Klapper D.G.;
RT "Isolation and characterization of a major cross-reactive grass group
RT I allergenic determinant.";
RL Mol. Immunol. 26:557-561(1989).
CC -1 SUBCELLULAR LOCATION: Secreted.
CC -1 DISEASE: CAUSES GRASS POLLEN ALLERGY.
CC -1 SIMILARITY: BELONGS TO THE LOL P I FAMILY OF ALLERGENS.
CC -1 SIMILARITY: CONTAINS 1 EXPANSIN-LIKE EGAS DOMAIN.
CC -1 SIMILARITY: CONTAINS 1 EXPANSIN-LIKE CBD DOMAIN.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 11, 2002, 14:50:30 ; Search time 5.73432 Seconds
(without alignments)
1499.000 Million cell updates/sec

Title: US-09-896-301-3

Perfect score: 1204
Sequence: 1 HMGPMINAHATFYXXGDAXX.....NVAPPTMSFGQTYTGCGFRY 222

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	217.5	18.1	263	1	MPPI_PHLPR
2	216	17.9	246	1	MPPI_CYNDA
3	215.5	17.9	263	1	MPPI_LOLPR
4	205.5	17.1	265	1	MPPI_HOLLA
5	202.5	16.8	269	1	MPPI_PHAQ
6	188.5	15.7	263	1	MPPI_ORYZA
7	178	14.8	191	1	MPPI_MAIZE
8	108.5	9.0	242	1	GUNS_TRIPE
9	92	7.6	333	1	CATB_MOUSE
10	84.5	7.0	922	1	PMPI_CHLUP
11	84	7.0	339	1	CATB_HUMAN
12	81	6.7	480	1	HRAL_MOUSE
13	79	6.6	164	1	CYPR_BLAG
14	78	6.5	649	1	KPCI_APLCA
15	78	6.5	1553	1	TPZA_CHICK
16	77.5	6.4	377	1	FL3H_HORVU
17	77	6.4	1115	1	GPCR_LYNST
18	77	6.4	1403	1	YGN1_YEAST
19	77	6.4	1451	1	VGL2_CVCAI
20	76.5	6.4	202	1	T4S4_HUMAN
21	76.5	6.4	305	1	Y758_ARCFU
22	76.5	6.4	335	1	CATB_BOVIN
23	76.5	6.4	340	1	CYSP_SCHMA
24	76	6.3	491	1	TYTR_CRIFA
25	75.5	6.3	213	1	GUNS_HUMAN
26	75	6.2	376	1	GUNK_FUSOX
27	75	6.2	1038	1	BMR2_HUMAN
28	74.5	6.2	504	1	G6PL_XANCI
29	74.5	6.2	546	1	ALRK_PSROL
30	74	6.1	493	1	CHT1_RHINI
31	74	6.1	526	1	TRM1_CAEEL
32	73.5	6.1	340	1	CATB_CHICK
33	73.5	6.1	363	1	VP43_NPVAC

34	73.5	6.1	491	1	TYTR_LEIDO	P39050 leishmania
35	73.5	6.1	1246	1	YMW2_CAEEL	P34504 caenorhabd
36	73	6.1	1038	1	BMR2_MOUSE	O35607 mus musculu
37	73	6.1	2491	1	MPRI_HUMAN	P11717 homo sapien
38	73	6.1	3010	1	POLG_HCVJA	P26662 h genome po
39	72.5	6.0	480	1	HRAL_HUMAN	O92743 homo sapien
40	72.5	6.0	1004	1	POIL_SCICO	O30277 sciara copr
41	72.5	6.0	319	1	ISL2_CHICK	P53410 gallus gall
42	71.5	5.9	308	1	HMB1_STRPU	P13545 strongyloce
43	71.5	5.9	502	1	G6PL_XYLEA	O99466 xylella fas
44	71.5	5.9	561	1	HEMA_IACKG	P12561 influenza a
45	71.5	5.9	561	1	HEMA_IAMIN	P12439 influenza a

ALIGNMENTS

RESULT 1

MPPI_PHLPR STANDARD: PRT: 263 AA.

AC P43213; 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DE 01-MAR-2002 (Rel. 41, Last annotation update)

GN pollen allergen Phl p 1 precursor (Phl p 1).

OS Phl p1.

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poaceae; Poaceae;

OC Poaceae; Phleum.

OX NCBI_TaxID=15957;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=pollen;

RX MEDLINE=95015525; PubMed=7930302;

RA Laffer S., Valenta R., Vrtala S., Susani M., van Ree R., Kraft D.,

RA Scheiner O., Duchene M.;

RT "Complementary DNA cloning of the major allergen Phl p 1 from timothy

RT grass (Phleum pratense): recombinant Phl p 1 inhibits IgE binding to

RT group 1 allergens from eight different grass species.";

RL J. Allergy Clin. Immunol. 94:689-698(1994).

CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- SIMILARITY: BELONGS TO THE LOL P I FAMILY OF ALLERGENS.

CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE EG45 DOMAIN.

CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE CBD DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@sib-sib.ch).

CC EMBL: X78813; CAA55390.1; -

DR HSSP: P43214; IMHO.

DR InterPro: IPRO00882; pollen.allergen.

DR Pfam: PF01357; Pollen.allergen; 1.

DR PRINTS: PR01225; EXPANSIN-FAMILY.

DR ProDom: PD002179; Pollen.allergen; 1.

DR PROSITE: PS50843; EXPANSIN-CBD; 1.

DR PROSITE: PS50842; EXPANSIN-EG45; 1.

KW Allergen; Glycoprotein; Signal; Multigene family.

FT SIGNAL 1 23 POTENTIAL.

FT CHAIN 24 263 POLLEN ALLERGEN PHL P 1.

FT DOMAIN 61 167 EXPANSIN-LIKE EG45.

FT DOMAIN 181 262 EXPANSIN-LIKE CBD.

FT CARBOHYD 32 32 N-LINKED (GLCNAC...) (POTENTIAL).

SO SQUONCE 263 AA; 28457 MW; 046BA249C17BC048 CRC64;

Query Match 18.1%; Score 217.5; DB 1; Length 263;

Best Local Similarity 28.2%; Pred. No. 1.2e-13;

Matches 66; Conservative 38; Mismatches 93; Indels 37; Gaps 11;


```

Db      83  CVDQPCWCLPG-STIVVATNECPNNALPNNAGWCNPLQHPDLAQVPQHTAQTKACI 141
      114  IPVYRRVRCKRSGGIRFTINGSHYFNLYLVTVGAGDVHVSVMKGSRTKQWLSRNG 173
      142  VPVAYRRIPCNKRGIRFTINGSHYFNLYLVTVGAGDVHVSVMKGSRTKQWLSRNG 201
      174  QNMOSNYLNGOSLSFVVTSDRSRVSPVAPPTWSPGOTY-TGGOF 220
      202  QNMGNANLNGOSLSFVVTSDGRTIVANNVASAGWFGOTYATGAOF 249

```

RESULT 15

```

Q9FUM3      PRELIMINARY:      PRT:      254 AA.
AC  Q9FUM3:
DT  01-MAR-2001 (Tremblrel. 16, Created)
DT  01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT  01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE  EXPANSIN 1.
GN  EXP1 OR EXP2.
OS  Prunus avium (Cherry), and
OC  Prunus cerasus.
OC  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC  eurosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.
OX  NCBI_TaxID=42229, 140311;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  SPECIES=P. avium;
RA  Wu Z., Wiersma P.A.;
RT  "Differential Expression of Expansin Genes Isolated from Sweet Cherry
RL  (Prunus avium L.) During Fruit Ripening.";
RN  Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
[2]
RP  SEQUENCE FROM N.A.
RC  SPECIES=P. cerasus; TISSUE=RIPENING FRUIT;
RA  Yoo S.-D., Gao Z., Cantini C., Loeschner W., van Nocker S.;
RT  "Coordinated expression of genes encoding expansins and other cell
RT  wall-modifying enzymes is associated with pectin-related changes in
RT  the cell wall during ripening of cherry (P. cerasus) fruit.";
RL  Submitted (Feb-2001) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AF297521; AAG13982.1; -
DR  EMBL; AF350937; AAK4846.1; -
DR  InterPro: IPR000882; Pollen_allergen.
DR  Pfam: PF01357; Pollen_allergen. 1.
DR  PRINTS: PR01225; EXPANSINFAMILY.
DR  Prodom: PD002179; Pollen_allergen. 1.
SQ  SEQUENCE 254 AA; 27278 MW; 953A7EB2491FD0E1 CRC64;

```

Query Match 73.9%; Score 889.5; DB 10; Length 254;
 Best local similarity 71.2%; Pred. No. 3.2e-82;
 Matches 161; Conservative 23; Mismatches 33; Indels 9; Gaps 2;

```

QY  3  GPVINAHATFFXXGDAXXTMGAGCYGNYLSQSGLETAALSTALFPDGLSCGACXELMC 62
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db  30  GMEGAHATFFYGGDASTMGAGCYGNYLSQSGLETAALSTALFPDGLSCGACXELMC 89
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY  63  VNDPQWCIKGRSIVTATNFCCP-----GCADPPNHFDPISOPLEYEKIALYKSGIT 114
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db  90  NNDPRWCPRG-SIVTATNFCCPFPFAOSNDNGWCNPPLOHFDIAEPALQIAQYRAGIV 148
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY  115  PVMTRRVACRKSIGIRFTINGSHYFNLYLVTVGAGDVHVSVMKGSRTKQWLSRNGQ 174
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db  149  PVTRFRVPCMKKGGIRFTINGSHYFNLYLVTVGAGDVHVSVMKGSRTKQWLSRNGQ 208
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY  175  NMOGSNYLNGOSLSFVVTSDRSRVSPVAPPTWSPGOTY-TGGOF 220
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db  209  NMOGSNYLNGOSLSFVVTSDGRTIVANNVASAGWFGOTYATGAOF 254
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

```

Search completed: October 11, 2002, 14:59:40
 Job time : 20.169 secs

QY 5 WINAHATFYXXGDAAXXTMGACGYGNLYSGYGLTALSTALFDGLSCGACXELMCVN 64
 DB 25 WVNNAHATFYGGSDASGTMGACGCGNLYSGYGTNTALSTALFNNGLSCGACFEIKCS 84
 QY 65 DPQWICIGRSIVTATNFCPP-----GCACDPNNHFDLSOPIYEKIALYKSGIT 116
 DB 85 DGACMLPG-ATIVTATNFCPPNNALPNNACGMCNPLPHFDLSQPVFQRIAYKAGVVP 143
 QY 117 MYRRVRCRSGCIRFTINGHSYFNLVLTNNGAGDVHSVMKSGRTKQMLSMNMGON 176
 DB 144 SYRRVPCRRGCIKFTINGHSYFNLVLTNNGAGDVHSAVAGSRTRKQMSRMNMGON 203
 QY 177 OSNLYNGQSLSFVTTSDRSVSVFNVAAPTMSFGQTYTGQGR 221
 DB 204 OSNNLNGQALSFVKTASDGRVTSNNIAPASMSFGQTFGRQPR 248

RESULT 12
 Q93XPI
 ID 093XPI PRELIMINARY; PRT; 249 AA.
 AC 093XPI
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE EXPANSIN.
 GN EXP3.
 OS Prunus cerasus.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC euroids I; Rosales; Rosaceae; Amygdaloideae; Prunus.
 OX NCBI_TaxID=140311;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=FRUIT ABSCISSION ZONE;
 RA Yoo S.-D., van Nocker S.;
 RL Submitted (Feb-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF350938; AAR4847.1;
 SQ SEQUENCE 249 AA; 26397 MW; C65556B038A99D0 CRC64;

Query Match 74.5%; Score 896.5; DB 10; Length 249;
 Best Local Similarity 71.7%; Pred. No. 6e-83;
 Matches 162; Conservative 24; Mismatches 31; Indels 9; Gaps 2;
 QY 3 GPMINAHATFYXXGDAAXXTMGACGYGNLYSGYGLTALSTALFDGLSCGACXELMC 62
 DB 25 GGMINAHATFYGGSDASGTMGACGCGNLYSGYGTNTALSTALFNNGLSCGACFEIKC 84
 QY 63 VNDPQWICIGRSIVTATNFCPP-----GCACDPNNHFDLSOPIYEKIALYKSGIT 114
 DB 85 ASDPKWMLPG-SIVTATNFCPPNNALPNNACGMCNPLPHFDLSQPVFQRIAYKAGV 143
 QY 115 PVMRVRRCRSGCIRFTINGHSYFNLVLTNNGAGDVHSVMKSGRTKQMLSMNMGON 174
 DB 144 PVSRRVPCRRGCIKFTINGHSYFNLVLTNNGAGDVHSAVAGSRTRKQMSRMNMGON 203
 QY 175 NMOSNLYNGQSLSFVTTSDRSVSVFNVAAPTMSFGQTYTGQGR 220
 DB 204 NMOSNLYNGQSLSFVKTASDGRVTSNNIAPASMSFGQTFGRQPR 249

RESULT 13
 O82093
 ID 082093 PRELIMINARY; PRT; 254 AA.
 AC 082093
 DT 01-NOV-1998 (Tremblrel. 08, Created)
 DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE EXPANSIN.
 GN PA-Exp1.
 OS Prunus armeniaca (Apicot).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC euroids I; Rosales; Rosaceae; Amygdaloideae; Prunus.
 OX NCBI_TaxID=36596;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERGERON; TISSUE=MESOCARP PLUS EXOCARP;
 RA Mbeugle-A-Mbeugle D., Gomez R.-M., Fils-Lycaon B.;
 RT "Molecular cloning and nucleotide sequence of expansin 1 (PA-Exp1)
 from apricot fruit."
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U93167; AAC33529.1;
 DR InterPro: IPR000882; Pollen_allergen.
 DR Pfam: PF01357; Pollen_allergen.1.
 DR PRINTS: PR01225; EXPANSINFAMLY.
 DR Prodom: PD002179; Pollen_allergen.1.
 SQ SEQUENCE 254 AA; 27264 MW; 88068D75932FDD0E1 CRC64;

Query Match 74.1%; Score 892.5; DB 10; Length 254;
 Best Local Similarity 71.7%; Pred. No. 1.6e-82;
 Matches 162; Conservative 22; Mismatches 33; Indels 9; Gaps 2;
 QY 3 GPMINAHATFYXXGDAAXXTMGACGYGNLYSGYGLTALSTALFDGLSCGACXELMC 62
 DB 30 GGMINAHATFYGGSDASGTMGACGCGNLYSGYGTNTALSTALFNNGLSCGACFEIKC 89
 QY 63 VNDPQWICIGRSIVTATNFCPP-----GCACDPNNHFDLSOPIYEKIALYKSGIT 114
 DB 90 NNDPQWICIGRSIVTATNFCPP-----GCACDPNNHFDLSOPIYEKIALYKSGIT 148
 QY 115 PVMRVRRCRSGCIRFTINGHSYFNLVLTNNGAGDVHSVMKSGRTKQMLSMNMGON 174
 DB 149 PVSRRVPCRRGCIKFTINGHSYFNLVLTNNGAGDVHSAVAGSRTRKQMSRMNMGON 208
 QY 175 NMOSNLYNGQSLSFVTTSDRSVSVFNVAAPTMSFGQTYTGQGR 220
 DB 209 NMOSNLYNGQSLSFVKTASDGRVTSNNIAPASMSFGQTFGRQPR 254

RESULT 14
 O9M515
 ID 09M515 PRELIMINARY; PRT; 249 AA.
 AC 09M515
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE ALPHA-EXPANSIN 1.
 OS Triphysaria versicolor.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Lamiales; Orobanchaceae; Triphysaria.
 OX NCBI_TaxID=64093;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=ROOTS TREATED WITH DMSO;
 RX MEDLINE=21186069; PubMed=11290422;
 RA Wobbel R.L., Yoder J.I.;
 RT "Differential RNA expression of alpha-expansin gene family members in
 the parasitic angiosperm Triphysaria versicolor (Scrophulariaceae).";
 RL Gene 266; 85-93(2001).
 DR EMBL: AF230278; AAF32411.1;
 DR InterPro: IPR000882; Pollen_allergen.
 DR PRINTS: PR01225; EXPANSINFAMLY.
 DR Prodom: PD002179; Pollen_allergen.1.
 SQ SEQUENCE 249 AA; 26116 MW; DC976FEE376794FF CRC64;

Query Match 74.1%; Score 892; DB 10; Length 249;
 Best Local Similarity 71.5%; Pred. No. 1.7e-82;
 Matches 163; Conservative 23; Mismatches 32; Indels 10; Gaps 3;
 QY 2 GPMINAHATFYXXGDAAXXTMGACGYGNLYSGYGLTALSTALFDGLSCGACXELMC 61
 DB 23 VGMINAHATFYGGSDASGTMGACGCGNLYSGYGTNTALSTALFNNGLSCGACFEIKC 82
 QY 62 VNDPQWICIGRSIVTATNFCPP-----GCACDPNNHFDLSOPIYEKIALYKSGIT 113

DT	01-NOV-1996 (TREMBlrel). 01, Last sequence update)
DT	01-DEC-2001 (TREMBlrel). 19, Last annotation update)
DE	EXPANSIN AT-EXPI (FRAGMENT).
GN	AT-EXPI.
OS	Arabidopsis thaliana (Mouse-ear cress).
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsis.
NC	NCBI_TaxID=3702;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=96016146; PubMed=7568110;
RA	Shcherban T.Y., Shi J., Durechko D.M., Gullitnan M.J.,
RA	McGueen-Mason S.J., Shieh M., Cosgrove D.J.;
RT	"Molecular cloning and sequence analysis of expansins--a highly
RT	conserved, multigene family of proteins that mediate cell wall
RT	extension in plants."
RL	Proc. Natl. Acad. Sci. U.S.A. 92:9245-9249(1995).
DR	EMBL; U30476; AAB38070.1; -
DR	InterPro: IPR000882; Pollen_allergen.
DR	Pfam: PF01357; Pollen_allergen; 1.
DR	PRINTS: PR01225; EXPANSINFAMILY.
DR	Prodom: PD002179; Pollen_allergen; 1.
FT	NON_TER
SO	SEQUENCE 237 AA; 25155 MW; 0E5A2CB9C29A3682 CRC64;
Query Match	75.5%; Score 909.5; DB 10; Length 237;
Best Local Similarity	73.1%; Pred No. 2,7e-84;
Matches 166; Conservative 19; Mismatches 33; Indels 9; Gaps 2;	
QY	3 GPMINAHATFTYXXGDXAXXTMGACGAGYGNLYSGQYGLFTALSTALFPDGLSGCAGXELMC 62
DB	12 GGVWVNHAFITFGGGDASGTMGACGAGYGNLYSGQYGNTPALSTALFPNGLSGCAFELRC 71
QY	63 VNDPWCIGKRSIVATATFCPP-----GCACDPRPHNHDLSDPIYEXLIALYKSGTI 114
DB	72 QNDGKWCPLRG-SIVYATATATFCPPNNALPRNMGAGMCNPRQDHNHLSQVDFORIDAYRAGIV 130
QY	115 PVYKRRVRCKRBSGIGRFTTNGHSYTNLVLYTVNGAGDVHSYKSGSTTKQMLSRMNGQ 174
DB	131 PVAYRRVPCVRRGGIGRFTTNGHSYTNLVLYTVNGAGDVHSYKSGSTTKQMLSRMNGQ 190
QY	175 NMOSNLYNGQSLSFVYTTSDRRSVYFNVAPPMPSFGQYTTGGQFR 221
DB	191 NMOSNLYNGQSLSFVYTTSDQYITVSNVAMNAGKSGQYTTGQMR 237

RESULT 9	
Q9ZP36	
ID	Q9ZP36
AC	Q9ZP36; PRELIMINARY;
DT	01-MAY-1999 (TREMBLrel. 10, Created)
DT	01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE	ALPHA-EXPANSIN PRECURSOR.
GN	NT-EXP4.
OS	Nicotiana tabacum (Common tobacco).
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX	NCBI_TaxID=4097;
RP	[1]
RN	SEQUENCE FROM N.A.
RC	STRAIN=CV. BY2.
RX	MEDLINE=99026297; PubMed=9808735;
RA	Link B.M., Cosgrove D.J.;
RT	"Acidic growth"

RT bright yellow response and alpha-expansins in suspension cultures of
RL bright yellow 2 tobacco.*;
PL Plant Physiol. 118:907-916(1998).
DR EMBL; AF049353; AAC96080.1; -
DR InterPro; IPR000882; Pollen.allergen.
DR Pfam; PF01357; Pollen.allergen.1.
DR PRINTS; PRO1225; EXPANSNFMALY.

DR Pfam; PF01357; Pollen_allergen: 1.
 DR PRINTS; PRO1225; EXPANSNFAMILY.
 DR Prodom; PD002179; Pollen_allergen: 1.
 SO SEQUENCE 255 AA; 27611 MW; 7580595A30DC414B CRC64;

Query Match 99.3%; Score 1195; DB 10; Length 255;
 Best Local Similarity 97.7%; Pred. No. 3.2e-113;
 Matches 217; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 HMGPIINAHATFYXXGDAAXXTMGACGYNLYSGYGLTALSTALFDGSLSCGACXEL 60
 DB 34 HMGPIINAHATFYXXGDAAXXTMGACGYNLYSGYGLTALSTALFDGSLSCGACXEL 93
 QY 61 MCVNDPQMCIRGSIIVATATNFCPPGACDPPNHHFELSOPRYEKIALYKSGIIPVWYRR 120
 DB 94 MCVNDPQMCIRGSIIVATATNFCPPGACDPPNHHFELSOPRYEKIALYKSGIIPVWYRR 153
 QY 121 VRCRSGGIRFTINGHSYFNLVLTNNGAGADVHSVSKSGRTKQMLSRMGMQMSNS 180
 DB 154 VRCRSGGIRFTINGHSYFNLVLTNNGAGADVHSVSKSGRTKQMLSRMGMQMSNS 213
 QY 181 YLNGSLSFVVTSDRRSVSFNVAAPPTWSFGQTYTGQGFRR 222
 DB 214 YLNGSLSFVVTSDRRSVSFNVAAPPTWSFGQTYTGQGFRR 255

RESULT 2

Q9FNT1 PRELIMINARY; PRT; 245 AA.

ID 09FNT1
 AC 09FNT1
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE EXPANSIN.
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eucosids I; Fabales; Fabaceae; Papilionoideae; Ciceraceae; Cicer.
 NCBI_TaxID=3827;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. CASTELLANA; TISSUE=ETIOLATED EPICOTYLS;
 RA Dopico B., Sanchez M.A., Labrador E.;
 RT "An expansin is expressed in chickpea epicotyls."
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ291816; CAC19183.1;
 DR InterPro; IPR000882; Pollen_allergen.
 DR PRINTS; PRO1225; EXPANSNFAMILY.
 DR Prodom; PD002179; Pollen_allergen: 1.
 SO SEQUENCE 245 AA; 26469 MW; 0E6BF6C835A38923 CRC64;

Query Match 83.3%; Score 1003; DB 10; Length 245;
 Best Local Similarity 82.6%; Pred. No. 9.2e-94;
 Matches 180; Conservative 13; Mismatches 25; Indels 0; Gaps 0;

QY 5 WINAHATFYXXGDAAXXTMGACGYNLYSGYGLTALSTALFDGSLSCGACXELMCVN 64
 DB 28 WINAHATFYXXGDAAXXTMGACGYNLYSGYGLTALSTALFDGSLSCGACXELMCVN 87
 QY 65 DPOMCIRGSIIVATATNFCPPGACDPPNHHFELSOPRYEKIALYKSGIIPVWYRR 124
 DB 88 DPOMCIRGSIIVATATNFCPPGACDPPNHHFELSOPRYEKIALYKSGIIPVWYRR 147
 QY 125 RSGGIRFTINGHSYFNLVLTNNGAGADVHSVSKSGRTKQMLSRMGMQMSNSYLNG 184
 DB 148 RSGGIRFTINGHSYFNLVLTNNGAGADVHSVSKSGRTKQMLSRMGMQMSNSYLNG 207
 QY 185 OSLSFVVTSDRRSVSFNVAAPPTWSFGQTYTGQGFRR 222
 DB 208 OSLSFVVTSDRRSVSFNVAAPPTWSFGQTYTGQGFRR 245

RESULT 3

Q92P31 PRELIMINARY; PRT; 239 AA.

ID Q92P31
 AC Q92P31
 DT 01-MAY-1999 (Tremblrel. 10, Created)
 DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE EXPANSIN PRECURSOR.
 GN EXP5.
 OS Lycopersicon esculentum (Tomato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 CC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 NCBI_TaxID=4081;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. 91347; TISSUE=FRUIT;
 RX MEDLINE=99178803; PubMed=10080718;
 RA Brumwell D.A., Harpster M.H., Dunsnault P.;
 RT "Differential expression of expansin gene family members during growth
 and ripening of tomato fruit."
 RL Plant Mol. Biol. 39:161-169(1999).
 DR EMBL; AF059489; AM13633.1;
 DR InterPro; IPR000882; Pollen_allergen.
 DR Pfam; PF01357; Pollen_allergen: 1.
 DR PRINTS; PRO1225; EXPANSNFAMILY.
 DR Prodom; PD002179; Pollen_allergen: 1.
 DR Signal.
 FT SIGNAL.
 FT CHAIN 21 239
 SQ SEQUENCE 239 AA; 25606 MW; 1C43BF3A1021788C CRC64;

Query Match 77.5%; Score 933.5; DB 10; Length 239;
 Best Local Similarity 77.2%; Pred. No. 1e-86;
 Matches 169; Conservative 20; Mismatches 29; Indels 1; Gaps 1;

QY 3 GPINAHATFYXXGDAAXXTMGACGYNLYSGYGLTALSTALFDGSLSCGACXELMC 62
 DB 22 GPINAHATFYXXGDAAXXTMGACGYNLYSGYGLTALSTALFDGSLSCGACXELMC 81
 QY 63 VNDPQMCIRGSIIVATATNFCPPGACDPPNHHFELSOPRYEKIALYKSGIIPVWYRR 122
 DB 82 VNDPQMCIRGSIIVATATNFCPPGACDPPNHHFELSOPRYEKIALYKSGIIPVWYRR 140
 QY 123 CRSGGIRFTINGHSYFNLVLTNNGAGADVHSVSKSGRTKQMLSRMGMQMSNSYL 182
 DB 141 CRSGGIRFTINGHSYFNLVLTNNGAGADVHSVSKSGRTKQMLSRMGMQMSNSYL 200
 QY 183 NGSLSFVVTSDRRSVSFNVAAPPTWSFGQTYTGQGFRR 221
 DB 201 NGSLSFVVTSDRRSVSFNVAAPPTWSFGQTYTGQGFRR 239

RESULT 4

Q9FUM2 PRELIMINARY; PRT; 252 AA.

ID Q9FUM2
 AC Q9FUM2
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE EXPANSIN 2.
 GN EXP2.
 OS Prunus avium (Cherry).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eucosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.
 NCBI_TaxID=42229;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wu Z., Wiersma P.A.;
 RT "Differential Expression of Expansin Genes Isolated from Sweet Cherry
 (Prunus avium L.) during Fruit Ripening".
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF297522; AM13983.1;

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 11, 2002, 14:51:00 : Search time 19.169 Seconds

(without alignments)
2003.488 Million cell updates/sec

Title: US-09-896-301-3
Perfect score: 1204

Sequence: 1 HMGPMWLNATFFXXGDAXX.....NVAPPTWSPGQTYTGQFRY 222

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: SPREMBL_19:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mhc:*
9: sp_organelle:*
10: sp_phage:*
11: sp_plant:*
12: sp_rhodent:*
13: sp_virus:*
14: sp_vertebrate:*
15: sp_unclassified:*
16: sp_virus:*
17: sp_bacteria:*
17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1195	99.3	255	10	Q38864
2	1003	83.3	245	10	Q9FNT1
3	933.5	77.5	239	10	Q92P31
4	931.5	77.4	252	10	Q9FUM2
5	929.5	77.2	247	10	Q9M517
6	928	77.1	242	10	Q9LIB1
7	927.5	77.0	249	10	Q9LDR9
8	909.5	75.5	237	10	Q38863
9	909.5	75.5	249	10	Q92P36
10	909.5	75.5	250	10	Q9C554
11	907.5	75.4	248	10	Q80622
12	896.5	74.5	249	10	Q93XP1
13	892.5	74.1	254	10	Q82093
14	892	74.1	249	10	Q9M515
15	889.5	73.9	254	10	Q9FUM3
16	886.5	73.6	245	10	Q9LIB2

17	874.5	72.6	260	10	Q9FNT0	Q9FNT0
18	873.5	72.5	253	10	Q93XP2	Q93XP2
19	871.5	72.4	253	10	Q9SMT1	Q9SMT1
20	871.5	72.4	255	10	Q9FMA0	Q9FMA0
21	870.5	72.3	232	10	Q93493	Q93493
22	870.5	72.3	253	10	Q9SWY1	Q9SWY1
23	867	72.0	252	10	Q9FS30	Q9FS30
24	865	71.8	252	10	Q81133	Q81133
25	864.5	71.8	232	10	Q93495	Q93495
26	863.5	71.7	232	10	Q93492	Q93492
27	861.5	71.6	253	10	Q9SMD4	Q9SMD4
28	857.5	71.2	250	10	Q93625	Q93625
29	856.5	71.1	232	10	Q93494	Q93494
30	856.5	71.1	250	10	Q9FV93	Q9FV93
31	855.5	71.1	262	10	Q80932	Q80932
32	852.5	70.8	247	10	Q82625	Q82625
33	852.5	70.8	258	10	Q49194	Q49194
34	850.5	70.6	249	10	Q92P35	Q92P35
35	845.5	70.2	258	10	Q9L299	Q9L299
36	842	69.9	253	10	Q22874	Q22874
37	832.5	69.1	257	10	Q48818	Q48818
38	832.5	69.1	258	10	Q41043	Q41043
39	825.5	68.6	260	10	Q9M259	Q9M259
40	825.5	68.6	276	10	Q94KT6	Q94KT6
41	824	68.4	255	10	Q9FIC5	Q9FIC5
42	823	68.4	255	10	Q38866	Q38866
43	820.5	68.1	251	10	Q946J1	Q946J1
44	816.5	67.8	257	10	Q9XG16	Q9XG16
45	815.5	67.7	246	10	P93442	P93442

ALIGNMENTS

RESULT 1

Q38864 PRELIMINARY: PRT; 255 AA.

AC Q38864;
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE EXPANSIN ATEX5 (EXPANSIN-LIKE PROTEIN).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96016146; PubMed=7568110;
RA Shcherban T.Y., Shi J., Durachko D.M., Gullifman M.J.,
RA McQueen-Mason S.J., Shieh M., Cosgrove D.J.;
RT "Molecular cloning and sequence analysis of expansin--a highly
RT conserved, multigene family of proteins that mediate cell wall
RT extension in plants.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:9245-9249(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
RN Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RA Nakamura Y.;
RX MEDLINE=20277480; PubMed=10819329;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. 1. Sequence
RT features of the regions of 4,504,864 bp covered by sixty p1 and TAC
RT clones.";
RT DNA Res. 7:131-135(2000).
DR EMBL; U30478; AAB38071.1; -;
DR EMBL; AB025615; BAA95756.1; -;
DR InterPro; IPR000882; Pollen_allergen.

PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 10-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149930.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 22-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0158029.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.

PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 88.6%; Score 115; DB 21; Length 281;
 Best Local Similarity 87.1%; Pred No. 17e-103;
 Matches 196; Conservative 13; Mismatches 16; Indels 0; Gaps 0;

QY 4 GWYNAHATFYGGDASGTMGACGYNLXSQYGTNTALSTALFNNGISGCAFEIRQ 63
 Db |||||||||
 Db 57 GWYNAHATFYGGDASGTMGACGYNLXSQYGTNTALSTALFNNGISGCAFEIRQ 116
 QY 64 NDKKWCIPGSTIVTATNFCPPNNALPNNAGKWCNPPQOHFDLSQPVFORIAQYRAGIVPV 123
 Db :|||
 QY 117 SDGAWCLPGAIIVTATNFCPPNNALPNNAGKWCNPPQOHFDLSQPVFORIAQYRAGIVPV 176
 Db :|||||
 QY 124 AYRRVPCVRGGIRFTINGHSYFNVLITNVGAGDVHSAMVKSRTGQAMSRMGMW 183
 Db :|||||
 QY 177 SYRRVPCMRGGIRFTINGHSYFNVLITNVGAGDVHSAMVKSRTGQAMSRMGMW 236
 QY 184 QSNVYLNQSLSPFKVTSDDQIVSNXNAGNSFGOTFGAHR 228
 Db |||:|||||
 Db 237 QSNVYLNQSLSPFKVTSDDQIVSNXNAGNSFGOTFGAHR 281

Search completed: October 11, 2002, 14:56:41
 Job time : 27.2399 secs

PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161923.
PR 29-OCT-1999; 99US-0162142.

Query Match 88.6% Score 115; DB 21; Length 253;
Best Local Similarity 87.1% Pred. No. 1,4e-103;
Matches 196; Conservative 13; Mismatches 16; Indels 0; Gaps 0;

QY 4 GVNNAHATFYGGDASGTMGACGYNLYSQYGTNTALSTALFENNLSCGACFEIRCO 63
DB 29 GVNNAHATFYGGDASGTMGACGYNLYSQYGTNTALSTALFENNLSCGACFEIRCO 88
QY 64 NDGKCLGSGIVTATNCPNNALPNNAGWCNPPQOHFPLSQPVQRIQYRAGIYV 123
DB 89 SDGAWCLPGLIIVTATNCPNNALPNNAGWCNPPQOHFPLSQPVQRIQYRAGIYV 148
QY 124 AYRRVPCVRRGIRFTINGHSTFNLITNNGAGDVHSAAVWKSRTGQMSRNMGONW 183
DB 149 SYRRVPCVRRGIRFTINGHSTFNLITNNGAGDVHSAAVWKSRTGQMSRNMGONW 208
QY 184 QSNVYLCQSLSEKVTTSDDQTIYSNNXANAGWSEFGOTFTGAHVR 228
DB 209 QSNVYLCQSLSEKVTTSDDQTIYSNNXANAGWSEFGOTFTGAHVR 253

RESULT 15

AAG43341
ID AAG43341 standard; Protein: 281 AA.

AC AAG43341;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 54161.

KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0121860.

PR 09-MAR-1999; 99US-0122548.

PR 23-MAR-1999; 99US-0123788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0128845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135529.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140891.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0143297.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144684.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.

PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141387.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.

PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159337.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.


```

PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147202.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161922.

PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match
Best Local Similarity 88.6%; Score 1115; DB 21; Length 253;
Matches 196; Conservative 13; Mismatches 16; Indels 0; Gaps 0;

QY 4 GWNNAHATFYGGDASGTMGACGYNLVSOGYGTPTALSTALFNNGLSCGACFEIKCO 63
|||||
DB 29 GWNNAHATFYGGSDASGTMGACGYNLVSOGYGTPTALSTALFNNGLSCGACFEIKCO 88
|||||
QY 64 NDKKWLPGSTIVYATNFCPPNNALPNNAGWCNPPQOHFDLSQPYFOKIAOTRAGIVY 123
|||||
DB 89 SDGAWCLPGAIIVTATNFCPPNNALPNNAGWCNPPHFDLSQPYFOKIAOTRAGIVY 148
|||||
QY 124 AYRRVPCVRRGIRFTINGSYFNLVLTINVGAGDVHSAMVGSRTGMQAMSRMGOW 183
|||||
DB 149 SYRRVPCVRRGIRFTINGSYFNLVLTINVGAGDVHSAMVGSRTGMQAMSRMGOW 208
|||||
QY 184 QSNNSYLGOSLSFKVTTSDQRTIVSNXNAGWSFGOTFTGAHVR 228
|||||
DB 209 QSNNLNGALSFKVTASDGRVTVSNNTIAPASWSFGOTFTGRQFR 253
|||||

RESULT 14
AAG43342
ID AAG43342 standard; Protein; 253 AA.
XX
AC AAG43342;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 54162.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129645.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 11-MAY-1999; 99US-0134258.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.

```

	Matches	196:	Conservative	13:	Mismatches	16:	Indels	0:	Gaps	0:
QY	4	GVVNAHATFYGGDASGTMGACGIGTGLYSQGTGNTAALSTALPNNGLSCGACEIRCO								63
Db	27	GVVNAHATFYGGSDASGTMGACGIGTGLYSQGTGNTAALSTALPNNGLSCGACEIRCO								86
QY	64	NDGKMCLEPSIVVATNFECPPNNALPNNAGGRCNPPOHFEEDLSOPVFORIAQYRAGIYPV								123
Db	87	NDGKMCLEPSIVVATNFECPPNNALPNNAGGRCNPPOHFEEDLSOPVFORIAQYRAGIYPV								146
QY	124	AYRVPVCVRREGIRFTINGHSYFNLVLITNVGAGDVHSAMVKGSRGTGQAMSRNWGMW								183
Db	147	SYRRVPCMRGRGIRFTINGHSYFNLVLITNVGAGDVHSVAVKGSRTRMQMSRWGMW								206
QY	184	QSNSTLNGOSLSFKYTFSDGCTIVSNKNXNAAGMSFGQFTTGAVHR								228
Db	207	QSNSTLNGOALSFKYVTSAGRTVSNNTIAPASWSFGQFTTGQPR								251
RESULT 13										
XX	AC	AAG23851								
XX	AC	AAG23851 standard; Protein: 253 AA.								
XX	DT	AAG23851;								
XX	DE	17-OCT-2000 (first entry)								
XX	DE	Arabidopsis thaliana protein fragment SEQ ID NO: 27310.								
KW	KW	Protein identification: signal transduction pathway; metabolic pathway;								
KM	KM	hybridisation assay; genetic mapping; gene expression control; promoter;								
XX	XX	termination sequence.								
XX	OS	Arabidopsis thaliana.								
XX	PN	EP1033405-A2.								
XX	PD	06-SEP-2000.								
XX	PF	25-FEB-2000; 2000EP-0301439.								
XX	PR	25-FEB-1999; 99US-0121825.								
XX	PR	05-MAR-1999; 99US-0123180.								
XX	PR	09-MAR-1999; 99US-0123548.								
XX	PR	23-MAR-1999; 99US-0125788.								
XX	PR	25-MAR-1999; 99US-0126264.								
XX	PR	29-MAR-1999; 99US-0126785.								
XX	PR	01-APR-1999; 99US-0127462.								
XX	PR	06-APR-1999; 99US-0128234.								
XX	PR	08-APR-1999; 99US-0128714.								
XX	PR	16-APR-1999; 99US-0129845.								
XX	PR	19-APR-1999; 99US-0130077.								
XX	PR	21-APR-1999; 99US-0130449.								
XX	PR	23-APR-1999; 99US-0130510.								
XX	PR	28-APR-1999; 99US-0130891.								
XX	PR	30-APR-1999; 99US-0131449.								
XX	PR	30-APR-1999; 99US-0132048.								
XX	PR	30-APR-1999; 99US-0132407.								
XX	PR	04-MAY-1999; 99US-0132484.								
XX	PR	05-MAY-1999; 99US-0132485.								
XX	PR	06-MAY-1999; 99US-0132486.								
XX	PR	06-MAY-1999; 99US-0132487.								
XX	PR	07-MAY-1999; 99US-0132863.								
XX	PR	11-MAY-1999; 99US-0134256.								
XX	PR	14-MAY-1999; 99US-0134218.								
XX	PR	14-MAY-1999; 99US-0134219.								
XX	PR	14-MAY-1999; 99US-0134221.								
XX	PR	14-MAY-1999; 99US-0134370.								
XX	PR	18-MAY-1999; 99US-0134768.								
XX	PR	19-MAY-1999; 99US-0134941.								
XX	PR	20-MAY-1999; 99US-0135124.								
XX	PR	21-MAY-1999; 99US-0135353.								

PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139452.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139751.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142802.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 21-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 21-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 22-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145216.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 03-AUG-1999; 99US-0146389.
 PR 04-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.

PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 22-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151338.
 PR 01-SEP-1999; 99US-0151530.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158359.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 18-OCT-1999; 99US-0159568.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161982.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 88.6%; Score 1115; DB 21; Length 251;
 Best Local Similarity 87.1%; Pred. No. 1.4e-103;

PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149175.
PR 20-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 07-SEP-1999; 99US-0151930.
PR 10-SEP-1999; 99US-0152363.
PR 13-SEP-1999; 99US-0153070.
PR 15-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154018.
PR 20-SEP-1999; 99US-0154039.
PR 22-SEP-1999; 99US-0154779.
PR 23-SEP-1999; 99US-0155139.
PR 24-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0155659.
PR 29-SEP-1999; 99US-0156458.
PR 04-OCT-1999; 99US-0156596.
PR 05-OCT-1999; 99US-0157117.
PR 06-OCT-1999; 99US-0157753.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160880.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
Query Match 88.6%; Score 1115; DB 21; Length 251;
Best Local Similarity 87.1%; Pred. No. 1.4e-103;
Matches 196; Conservative 13; Mismatches 16; Indels 0; Gaps 0;

OY 4 GWNAHATFEYGGSDASGTMGACGYNLYSOGYGTNTALSTALFLNNGLSCGACFEIRCO 63
DB 27 GWNAHATFEYGGSDASGTMGACGYNLYSOGYGTNTALSTALFLNNGLSCGACFEIRCO 86

OY 64 NDGKWCIPGSTIVTATNFCPPNNALPNNAGGMCNPPQOHFDLSQPVFORIAQYRAGTIVP 123
DB 87 SDGAWCLPGAIIVYATNFCPPNNALPNNAGGMCNPLPHFDLSQPVFORIAQYRAGTIVP 146
OY 124 AYRRVPCVRGRIFFETTINGSHYFNLYITVWGAAGVHSAVAGSRTRMQMSRRNGOWN 183
DB 147 SYRRVPCMRKGGRIFFETTINGSHYFNLYITVWGAAGVHSAVAGSRTRMQMSRRNGOWN 206
OY 184 QSNSTYLNQGSISFKVTTSDGQITVSNXNANAGMSFGQTFRTGAHR 228
DB 207 QSNNTLNGQALSPFKVTASDGRFTVSNNTIAPASMSFGQTFRTGRFR 251
RESULT 12
AAG43343
ID AAG43343 standard; Protein; 251 AA.
AC AAG43343;
XX
DT 18-OCT-2000 (first entry)
DE
XX Arabidopsis thaliana protein fragment SEQ ID NO: 54163.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
OS
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
PF
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 14-MAY-1999; 99US-0134376.
PR 18-MAY-1999; 99US-0134941.
PR 19-MAY-1999; 99US-0135124.
PR 20-MAY-1999; 99US-0135125.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.

PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145087.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.

PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 90.5%; Score 1139; DB 21; Length 280;
Best Local Similarity 89.0%; Pred. No. 6; 5e-106;
Matches 202; Conservative 12; Mismatches 13; Indels 0; Gaps 0;

OY 2 GGGWVAHAATFEYGGDASGTMGACGYGNLYSGYGTNTAALSTALFNNGLSGACFEIR 61
DB 54 GGGWVAHAATFEYGGDASGTMGACGYGNLYSGYGTNTAALSTALFNNGLSGACFEIR 113
OY 62 CONDGKWCPLPGSLIVYATNFCCPNNALPNNAGGWCNPPQOHFDLSQPVFORIAQYRAGIV 121
DB 114 CENDGKWCPLPGSLIVYATNFCCPNNALPNNAGGWCNPPLEHFDIAQPVFORIAQYRAGIV 173

```
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match      90.5%; Score 1139; DB 21; Length 259;
Best Local Similarity 89.0%; Pred. No. 5.9e-106;
Matches 202; Conservative 12; Mismatches 13; Indels 0; Gaps 0;
```

```
OY 2 GCGVNAHATRYGGGDSAGTGMGACGYNLXSQGYGTALSTALFNNGSCGACFEIR 61
    |||||
DB 33 GGGWNAHATRYGGGDSAGTGMGACGYNLXSQGYGTALSTALFNNGSCGACFEIR 92
    |||||
OY 62 CONDKMCLPSSIVYTANFCPPNNALPNNAGCNPPOQHFDLSOPVFORIAQYRAGIV 121
    |||||
DB 93 CENDCKMCLPSSIVYTANFCPPNNALPNNAGCNPPLLEHFDLAQYVFORIAQYRAGIV 152
    |||||
OY 122 PVAYRRVPCVARGIRFTINGHSYFNVLITNVGAGDVHSAMVKGSRFTQAMSRMNGQ 181
    |||||
DB 153 PVSRYRRVPCVARGIRFTINGHSYFNVLITNVGAGDVHSAMVKGSRFTQAMSRMNGQ 212
    |||||
OY 182 NMQSNSTYLNQALSFKYVTSDOQTIVSNKXANACHSFOQTFTGAHVR 228
```

```
DB 213 NMQSNSTYLNQALSFKYVTSDOQTIVSNKXANACHSFOQTFTGAHVR 259
    |||||
RESULT 10
ID AAG51645 standard; Protein; 280 AA.
AC AAG51645;
XX
XX 18-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 65566.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.
XX
XX Arabidopsis thaliana.
OS
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 22-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126284.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138840.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
```

AA51631
ID AA51631 standard; Protein; 259 AA.
XX
AC AA51631.
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 65549.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121925.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0148368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.

PR	18-JUN-1999:	99US-0139763.
PR	21-JUN-1999:	99US-0139817.
PR	22-JUN-1999:	99US-0139817.
PR	23-JUN-1999:	99US-0140353.
PR	23-JUN-1999:	99US-0140354.
PR	24-JUN-1999:	99US-0140695.
PR	28-JUN-1999:	99US-0140823.
PR	29-JUN-1999:	99US-0140991.
PR	30-JUN-1999:	99US-0141287.
PR	01-JUL-1999:	99US-0141842.
PR	01-JUL-1999:	99US-0142154.
PR	02-JUL-1999:	99US-0142055.
PR	06-JUL-1999:	99US-0142390.
PR	08-JUL-1999:	99US-0142803.
PR	09-JUL-1999:	99US-0142920.
PR	12-JUL-1999:	99US-0142977.
PR	13-JUL-1999:	99US-0143542.
PR	14-JUL-1999:	99US-0143624.
PR	15-JUL-1999:	99US-0144005.
PR	16-JUL-1999:	99US-0144085.
PR	16-JUL-1999:	99US-0144086.
PR	19-JUL-1999:	99US-0144325.
PR	19-JUL-1999:	99US-0144331.
PR	19-JUL-1999:	99US-0144332.
PR	19-JUL-1999:	99US-0144333.
PR	19-JUL-1999:	99US-0144334.
PR	19-JUL-1999:	99US-0144335.
PR	20-JUL-1999:	99US-0144352.
PR	20-JUL-1999:	99US-0144632.
PR	20-JUL-1999:	99US-0144884.
PR	21-JUL-1999:	99US-0144814.
PR	21-JUL-1999:	99US-0145086.
PR	21-JUL-1999:	99US-0145088.
PR	22-JUL-1999:	99US-0145085.
PR	22-JUL-1999:	99US-0145087.
PR	22-JUL-1999:	99US-0145089.
PR	22-JUL-1999:	99US-0145192.
PR	23-JUL-1999:	99US-0145145.
PR	23-JUL-1999:	99US-0145218.
PR	26-JUL-1999:	99US-0145224.
PR	27-JUL-1999:	99US-0145276.
PR	27-JUL-1999:	99US-0145913.
PR	27-JUL-1999:	99US-0145918.
PR	28-JUL-1999:	99US-0145919.
PR	02-AUG-1999:	99US-0146386.
PR	02-AUG-1999:	99US-0146388.
PR	02-AUG-1999:	99US-0146389.
PR	03-AUG-1999:	99US-0147038.
PR	04-AUG-1999:	99US-0147204.
PR	04-AUG-1999:	99US-0147302.
PR	05-AUG-1999:	99US-0147192.
PR	06-AUG-1999:	99US-0147260.
PR	06-AUG-1999:	99US-0147303.
PR	06-AUG-1999:	99US-0147416.
PR	09-AUG-1999:	99US-0147493.
PR	09-AUG-1999:	99US-0147935.
PR	10-AUG-1999:	99US-0148171.
PR	11-AUG-1999:	99US-0148319.
PR	12-AUG-1999:	99US-0148341.
PR	13-AUG-1999:	99US-0148565.
PR	13-AUG-1999:	99US-0148684.
PR	16-AUG-1999:	99US-0149388.
PR	17-AUG-1999:	99US-0149175.
PR	18-AUG-1999:	99US-0149426.
PR	20-AUG-1999:	99US-0149722.
PR	20-AUG-1999:	99US-0149929.
PR	23-AUG-1999:	99US-0149920.
PR	23-AUG-1999:	99US-0150566.
PR	26-AUG-1999:	99US-0150884.
PR	27-AUG-1999:	99US-0151065.
PR	27-AUG-1999:	99US-0151065.
PR	27-AUG-1999:	99US-0151066.
PR	27-AUG-1999:	99US-0151080.
PR	30-AUG-1999:	99US-0151303.
PR	31-AUG-1999:	99US-0151438.
PR	01-SEP-1999:	99US-0151930.
PR	07-SEP-1999:	99US-0152363.
PR	10-SEP-1999:	99US-0153070.
PR	13-SEP-1999:	99US-0153758.
PR	15-SEP-1999:	99US-0154018.
PR	16-SEP-1999:	99US-0154039.
PR	20-SEP-1999:	99US-0154779.
PR	22-SEP-1999:	99US-0155139.
PR	23-SEP-1999:	99US-0155486.
PR	24-SEP-1999:	99US-0155659.
PR	28-SEP-1999:	99US-0156458.
PR	29-SEP-1999:	99US-0156596.
PR	04-OCT-1999:	99US-0157117.
PR	05-OCT-1999:	99US-0157753.
PR	06-OCT-1999:	99US-0157865.
PR	07-OCT-1999:	99US-0158029.
PR	08-OCT-1999:	99US-0158232.
PR	12-OCT-1999:	99US-0158369.
PR	13-OCT-1999:	99US-0159293.
PR	13-OCT-1999:	99US-0159294.
PR	13-OCT-1999:	99US-0159295.
PR	14-OCT-1999:	99US-0159329.
PR	14-OCT-1999:	99US-0159330.
PR	14-OCT-1999:	99US-0159331.
PR	14-OCT-1999:	99US-0159637.
PR	14-OCT-1999:	99US-0159638.
PR	18-OCT-1999:	99US-0159584.
PR	21-OCT	

PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match

90.5%; Score 1139; DB 21; Length 249;

Best Local Similarity 89.0%; Pred. No. 5.6e-106; Matches 202; Conservative 12; Mismatches 13; Indels 0; Gaps 0;

OY 2 GGGVNAHATPYGGDASGTMGACGCGNLYSGYGTNTALSTALNNGLSCGACEIR 61
DB 23 GGGVNAHATPYGGDASGTMGACGCGNLYSGYGTNTALSTALNNGLSCGACEIR 82
OY 62 CONGKNCILPGSIYVTAFTNFCPPNNALPNNAGWCNPPPOHFDLSOPVFORIAOYRAGIV 121
DB 83 CENGGKNCILPGSIYVTAFTNFCPPNNALPNNAGWCNPPLEHFDLAQVFORIAOYRAGIV 142
OY 122 PVAYRVPVCRRGIRFTINGSHYFNVLITNVGAGDVHSAMVKGSRGTQWAMSRMWQ 181
DB 143 PVYRVPVCRRGIRFTINGSHYFNVLITNVGAGDVHSAMVKGSRGTQWAMSRMWQ 202
OY 182 NMOSNSTLNGOSLSFKYTTSDGOTIVSNNAAGWSGQFTTGAHVR 228
DB 203 NMOSNSTLNGOSLSFKYTTSDGRTVSEFNAPAGWSGQFTTGAQGR 249

RESULT 8
AAG06544
ID AAG06544 standard; Protein; 259 AA.
XX
AC AAG06544;
XX

DT 17-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 3355.
DE
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
OS
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123348.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 16-APR-1999; 99US-0128714.
PR 19-APR-1999; 99US-0129845.
PR 21-APR-1999; 99US-0130077.
PR 23-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 24-MAY-1999; 99US-0135533.
PR 25-MAY-1999; 99US-0135629.
PR 27-MAY-1999; 99US-0136021.
PR 28-MAY-1999; 99US-0136382.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.

hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.

Arabidopsis thaliana.

EPI033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.

PR 10-JUN-1999; 99US-0138847.

PR 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.

PR 17-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139750.

PR 18-JUN-1999; 99US-0139763.

PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139889.

PR 23-JUN-1999; 99US-0140353.

PR 23-JUN-1999; 99US-0140354.

PR 24-JUN-1999; 99US-0140695.

PR 28-JUN-1999; 99US-0140823.

PR 29-JUN-1999; 99US-0140991.

PR 30-JUN-1999; 99US-0141287.

PR 01-JUL-1999; 99US-0141842.

PR 01-JUL-1999; 99US-0142154.

PR 02-JUL-1999; 99US-0142055.

PR 06-JUL-1999; 99US-0142390.

PR 08-JUL-1999; 99US-0142803.

PR 09-JUL-1999; 99US-0142820.

PR 12-JUL-1999; 99US-0142877.

PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144085.

PR 16-JUL-1999; 99US-0144086.

PR 19-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.

PR 19-JUL-1999; 99US-0144332.

PR 19-JUL-1999; 99US-0144333.

PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144352.

PR 20-JUL-1999; 99US-0144632.

PR 20-JUL-1999; 99US-0144684.

PR 21-JUL-1999; 99US-0144814.

PR 21-JUL-1999; 99US-0145086.

PR 22-JUL-1999; 99US-0145088.

PR 22-JUL-1999; 99US-0145085.

PR 22-JUL-1999; 99US-0145087.

PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145192.

PR 23-JUL-1999; 99US-0145218.

PR 23-JUL-1999; 99US-0145224.

PR 25-JUL-1999; 99US-0145276.

PR 27-JUL-1999; 99US-0145813.

PR 27-JUL-1999; 99US-0145918.

PR 27-JUL-1999; 99US-0145919.

PR 28-JUL-1999; 99US-0145951.

PR 02-AUG-1999; 99US-0146386.

PR 02-AUG-1999; 99US-0146388.

PR 02-AUG-1999; 99US-0146389.

PR 03-AUG-1999; 99US-0147038.

PR 04-AUG-1999; 99US-0147204.

PR 04-AUG-1999; 99US-0147302.

PR 05-AUG-1999; 99US-0147192.

PR 05-AUG-1999; 99US-0147260.

PR 06-AUG-1999; 99US-0147303.

PR 06-AUG-1999; 99US-0147416.

PR 09-AUG-1999; 99US-0147493.

PR 09-AUG-1999; 99US-0147935.

PR 10-AUG-1999; 99US-0148171.

PR 11-AUG-1999; 99US-0148319.

PR 12-AUG-1999; 99US-0148341.

PR 13-AUG-1999; 99US-0148365.

PR 13-AUG-1999; 99US-0148684.

PR 16-AUG-1999; 99US-0148684.

PR 17-AUG-1999; 99US-0149175.

PR 18-AUG-1999; 99US-0149426.

PR 20-AUG-1999; 99US-0149722.

PR 20-AUG-1999; 99US-0149723.

PR 20-AUG-1999; 99US-0149729.

PR 23-AUG-1999; 99US-0149902.

PR 23-AUG-1999; 99US-0149930.

PR 25-AUG-1999; 99US-0150566.

PR 26-AUG-1999; 99US-0150884.

PR 27-AUG-1999; 99US-0151065.

PR 27-AUG-1999; 99US-0151066.

PR 27-AUG-1999; 99US-0151080.

PR 30-AUG-1999; 99US-0151303.

PR 31-AUG-1999; 99US-0151438.

PR 01-SEP-1999; 99US-0151930.

PR	01-JUL-1999;	99US-0142154.
PR	02-JUL-1999;	99US-0142035.
PR	06-JUL-1999;	99US-0142330.
PR	08-JUL-1999;	99US-0142803.
PR	09-JUL-1999;	99US-0142920.
PR	12-JUL-1999;	99US-0142977.
PR	13-JUL-1999;	99US-0143542.
PR	14-JUL-1999;	99US-0143624.
PR	15-JUL-1999;	99US-0144005.
PR	16-JUL-1999;	99US-0144085.
PR	16-JUL-1999;	99US-0144325.
PR	19-JUL-1999;	99US-0144331.
PR	19-JUL-1999;	99US-0144332.
PR	19-JUL-1999;	99US-0144333.
PR	19-JUL-1999;	99US-0144334.
PR	19-JUL-1999;	99US-0144335.
PR	20-JUL-1999;	99US-0144352.
PR	20-JUL-1999;	99US-0144632.
PR	20-JUL-1999;	99US-0144684.
PR	21-JUL-1999;	99US-0144814.
PR	21-JUL-1999;	99US-0145086.
PR	21-JUL-1999;	99US-0145088.
PR	22-JUL-1999;	99US-0145085.
PR	22-JUL-1999;	99US-0145087.
PR	22-JUL-1999;	99US-0145089.
PR	22-JUL-1999;	99US-0145192.
PR	23-JUL-1999;	99US-0145145.
PR	23-JUL-1999;	99US-0145218.
PR	23-JUL-1999;	99US-0145224.
PR	26-JUL-1999;	99US-0145276.
PR	27-JUL-1999;	99US-0145913.
PR	27-JUL-1999;	99US-0145918.
PR	27-JUL-1999;	99US-0145919.
PR	28-JUL-1999;	99US-0145951.
PR	02-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	02-AUG-1999;	99US-0146389.
PR	03-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	04-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147192.
PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149923.
PR	23-AUG-1999;	99US-0149920.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT	

PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160757.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 90.5%; Score 1139; DB 21; Length 249;
Best Local Similarity 89.0%; Pred. No. 5.6e-106;
Matches 202; Conservative 12; Mismatches 13; Indels 0; Gaps 0;

OY 2 GGGWNAATATYGGGADAGTGGACGACGYNLYSOGYGTATLALFNNGLSGACFEIR 61
DB 23 GGGWNAATATYGGGADAGTGGACGACGYNLYSOGYGTATLALFNNGLSGACFEIR 82
OY 62 CONDGKMLCPESIVYTATNFCPPNNALPNNAGWCNPPQOHFDLSOPVFORIAQYRAGIV 121
DB 83 CENCGKMLCPESIVYTATNFCPPNNALPNNAGWCNPPLEHFDLAQPFORIAQYRAGIV 142
OY 122 PVAYRRVPCVRGGIRFTINGHSYFNLVLTNVGAGDVHSAMVKSRTGMQAMSRMGO 181
DB 143 PVSYRRVPCVRGGIRFTINGHSYFNLVLTNVGAGDVHSAMVKSRTGMQAMSRMGO 202
OY 182 NMOSNYSYNGOSLSFKVTTSDGOTIVSNXNANAGMSFGOTFGAHR 228
DB 203 NMOSNYSYNGOSLSFKVTTSDGOTIVSNXNANAGMSFGOTFGAHR 249

RESULT 6
AAG51632
ID AAG51632 standard; Protein: 249 AA.

XX AAG51632;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 65550.

KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.

XX Arabidopsis thaliana.

XX

PN EF1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134376.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136382.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138647.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.

XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 25-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140685.
PR 28-JUN-1999; 99US-0140883.
PR 29-JUN-1999; 99US-0140931.
PR 30-JUN-1999; 99US-0141287.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142350.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.

PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151067.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.

PR	16-JUL-1999;	99US-0144086;	PR	08-OCT-1999;	99US-0158232;
PR	19-JUL-1999;	99US-0144325;	PR	12-OCT-1999;	99US-0158369;
PR	19-JUL-1999;	99US-0144331;	PR	13-OCT-1999;	99US-0159293;
PR	19-JUL-1999;	99US-0144332;	PR	13-OCT-1999;	99US-0159294;
PR	19-JUL-1999;	99US-0144333;	PR	13-OCT-1999;	99US-0159295;
PR	19-JUL-1999;	99US-0144334;	PR	14-OCT-1999;	99US-0159329;
PR	19-JUL-1999;	99US-0144335;	PR	14-OCT-1999;	99US-0159330;
PR	20-JUL-1999;	99US-0144352;	PR	14-OCT-1999;	99US-0159331;
PR	20-JUL-1999;	99US-0144632;	PR	14-OCT-1999;	99US-0159637;
PR	20-JUL-1999;	99US-0144684;	PR	14-OCT-1999;	99US-0159638;
PR	21-JUL-1999;	99US-0144814;	PR	18-OCT-1999;	99US-0159584;
PR	21-JUL-1999;	99US-0145086;	PR	21-OCT-1999;	99US-0160741;
PR	21-JUL-1999;	99US-0145088;	PR	21-OCT-1999;	99US-0160767;
PR	22-JUL-1999;	99US-0145085;	PR	21-OCT-1999;	99US-0160768;
PR	22-JUL-1999;	99US-0145087;	PR	21-OCT-1999;	99US-0160770;
PR	22-JUL-1999;	99US-0145089;	PR	21-OCT-1999;	99US-0160814;
PR	22-JUL-1999;	99US-0145192;	PR	21-OCT-1999;	99US-0160815;
PR	23-JUL-1999;	99US-0145145;	PR	22-OCT-1999;	99US-0160980;
PR	23-JUL-1999;	99US-0145218;	PR	22-OCT-1999;	99US-0160981;
PR	23-JUL-1999;	99US-0145224;	PR	22-OCT-1999;	99US-0160989;
PR	26-JUL-1999;	99US-0145276;	PR	25-OCT-1999;	99US-0161404;
PR	27-JUL-1999;	99US-0145913;	PR	25-OCT-1999;	99US-0161405;
PR	27-JUL-1999;	99US-0145918;	PR	25-OCT-1999;	99US-0161406;
PR	27-JUL-1999;	99US-0145919;	PR	26-OCT-1999;	99US-0161399;
PR	28-JUL-1999;	99US-0145951;	PR	26-OCT-1999;	99US-0161360;
PR	02-AUG-1999;	99US-0146386;	PR	26-OCT-1999;	99US-0161361;
PR	02-AUG-1999;	99US-0146388;	PR	28-OCT-1999;	99US-0161920;
PR	02-AUG-1999;	99US-0146389;	PR	28-OCT-1999;	99US-0161992;
PR	03-AUG-1999;	99US-0147038;	PR	28-OCT-1999;	99US-0161993;
PR	04-AUG-1999;	99US-0147204;	PR	29-OCT-1999;	99US-0162142;

Query Match	Score	DB 21;	Length
90.58;	1139;	21;	241;

Best Local Similarity	89.06	Pred. NO. 5.3e-100
Matches 202; Conservative	12;	Mismatches 13; Indels 0; Gaps 0;

2 GGGVNAHATFYGGDASGTMGACGYNLYSQGYGTNTALSTALFNNGLSGACFEIR 61

Db 15 GCGWNAHATFYGGGDASGTMGACGYNLYSQGYCTSTALSTALFNNGLSGSCFEIR 74

QY 62 CQNDGKWCCLPGSIVTATNFCCPPNNALPNNAGGWCNPPQOHFDLSQPVFORIAQYRAGIV 12

Db 75 CENDGKWC L P G S I V T A T N E C P P N N A L A N N G G W C N P L E H E D L A Q P V F Q R I A Q Y R A G I V 13

QY 122 PVAYRRVPCVRRGGIRETINGHSYFNLVLITNVGGAGDVHSAMVKGSRTGWQAMSRNWGQ 18

Db 135 PVS YRVP CRRGG IRT INGH SYNLV LITN VGGAGDVHSA IKGSR TVWQAMSRNWGQ 135

QY 182 NWQNSYLNQSLSEKVTTS DGQTI VSNNXANAGMSFGQTF TGAHVR 228

Db 195 NWQNSYLNQALSEKVTTS DGRVVSFNAAPAGWSYGQTFAGGQFR 241

AAG06545

XX

XX

XX

XX

KW hybridisation assay; genetic mapping; gene expression control; promoter

XX

XX

XX

XX

PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.

Query Match	90.5%;	Score 1139;	DB 21;	Length 241;
-------------	--------	-------------	--------	-------------

Best Local Similarity 89.0%; Pred. No. 5.3e-106;
Matches 202; Conservative 12; Mismatches 13; Indels 0; Gaps 0;

QY 2 CGGWNNAHATFYGGGDSCTMGACGCGNLYSGGCTNTAALSTALENNGLSGACFEIR 61

Db 15 GCGWINAHATFEYGGGDASGTMGACGCGNLYSQYGTSTALSTALENNGLSCGSCFEIR 74

0Y 62 CQNDGKWCPLGSIIVTATNFCPPNNALPNNAGCWCNPPQHQHFDLSQPVFEQRIAQYRAGIV 12

Db 75 CENDGKWCLEPGSIVTATNFCPPNNALANNNGWCNPLEHFDLAQPVFQIRIAQYRAGIV 13

```

122 PVAIRKVEPCYKRGSIKRTINGHSYFNLVLITINVGAGDVHSAMVKGSRTIGWQAMSRNWQ 18
||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

[illegible]

106 ATGCGACGT NCCAT CCGTCCTGGCCTGTTGCATAA TTA GAGTCAATTAA GCGGGG 341
||||| |: ||||| : | | | | |
||||| |: ||||| : | | | | |

RESULT 4
AAG51647

XX ID AAC3184 / standard; protein; 241 AA.

	AC	AMJ104 / i
	XX	

XX
XX
XX

(unclassified)

ambiguities +believe unclassified statement CTO TD NO CFF70 DE

protein identification. signal transduction pathway. metabolic pathway

KW hybridisation assay; genetic mapping; gene expression control; promoter termination sequence.

XX Arabidopsis thaliana.

AA
PN EP1033405-A2.

PD 06-SEP-2000.

PF -25-FEB-2000; 2000EP-0301439.
YY

PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180

PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.

PR 23-MAR-1999; 99US-0125788.

PR 23-MAR-1999; 99US-0125788.

99/CZT0-S056
23 MAR 1999
FN

PR	25-MAR-1999	9905-01267824
PR	25-MAR-1999	9905-01267825
PR	25-MAR-1999	9905-01267826
PR	01-APR-1999	9905-01274452
PR	06-APR-1999	9905-01281824
PR	08-APR-1999	9905-01287144
PR	16-APR-1999	9905-01329845
PR	19-APR-1999	9905-01300877
PR	21-APR-1999	9905-01304449
PR	23-APR-1999	9905-01303501
PR	23-APR-1999	9905-01308911
PR	28-APR-1999	9905-01311444
PR	30-APR-1999	9905-01312048
PR	30-APR-1999	9905-01324077
PR	04-MAY-1999	9905-01324445
PR	05-MAY-1999	9905-01324454
PR	06-MAY-1999	9905-01324467
PR	06-MAY-1999	9905-01324678
PR	07-MAY-1999	9905-01329863
PR	11-MAY-1999	9905-01342566
PR	14-MAY-1999	9905-01342118
PR	14-MAY-1999	9905-01342119
PR	14-MAY-1999	9905-01344221
PR	14-MAY-1999	9905-01343701
PR	18-MAY-1999	9905-01347668
PR	19-MAY-1999	9905-01349411
PR	20-MAY-1999	9905-01351124
PR	21-MAY-1999	9905-01353533
PR	24-MAY-1999	9905-01356629
PR	25-MAY-1999	9905-01360211
PR	27-MAY-1999	9905-01363892
PR	28-MAY-1999	9905-01367822
PR	01-JUN-1999	9905-01372222
PR	03-JUN-1999	9905-01375728
PR	04-JUN-1999	9905-01375052
PR	07-JUN-1999	9905-01377724
PR	08-JUN-1999	9905-01380924
PR	10-JUN-1999	9905-01385400
PR	10-JUN-1999	9905-01388847
PR	14-JUN-1999	9905-01391119
PR	16-JUN-1999	9905-01394452
PR	16-JUN-1999	9905-01394453
PR	17-JUN-1999	9905-01394452
PR	18-JUN-1999	9905-01394454
PR	18-JUN-1999	9905-01394455
PR	18-JUN-1999	9905-01394577
PR	18-JUN-1999	9905-01394578
PR	18-JUN-1999	9905-01394589
PR	18-JUN-1999	9905-01394600
PR	18-JUN-1999	9905-01394601
PR	18-JUN-1999	9905-01394602
PR	18-JUN-1999	9905-01394603
PR	18-JUN-1999	9905-01394604
PR	18-JUN-1999	9905-01394605
PR	18-JUN-1999	9905-01394606
PR	18-JUN-1999	9905-01394607
PR	21-JUN-1999	9905-01396817
PR	22-JUN-1999	9905-01396819
PR	22-JUN-1999	9905-01403553
PR	23-JUN-1999	9905-01403554
PR	24-JUN-1999	9905-01406655
PR	28-JUN-1999	9905-01408823
PR	28-JUN-1999	9905-01409911
PR	30-JUN-1999	9905-01411847
PR	01-JUL-1999	9905-01412822
PR	01-JUL-1999	9905-01421254
PR	02-JUL-1999	9905-01425055
PR	06-JUL-1999	9905-01422390
PR	08-JUL-1999	9905-01428200
PR	09-JUL-1999	9905-01429280
PR	12-JUL-1999	9905-01429377
PR	13-JUL-1999	9905-01435472
PR	14-JUL-1999	9905-01436324
PR	15-JUL-1999	9905-01440005
PR	16-JUL-1999	9905-01440805

PR	16-APR-1999;	9905-0129845;
PR	16-APR-1999;	9905-0130077;
PR	19-APR-1999;	9905-0130044;
PR	21-APR-1999;	9905-0130510;
PR	23-APR-1999;	9905-0130891;
PR	28-APR-1999;	9905-0131448;
PR	30-APR-1999;	9905-0132048;
PR	30-APR-1999;	9905-0132407;
PR	04-MAY-1999;	9905-0132484;
PR	05-MAY-1999;	9905-0132485;
PR	06-MAY-1999;	9905-0132486;
PR	06-MAY-1999;	9905-0132487;
PR	07-MAY-1999;	9905-0132863;
PR	11-MAY-1999;	9905-0134256;
PR	14-MAY-1999;	9905-0134218;
PR	14-MAY-1999;	9905-0134221;
PR	14-MAY-1999;	9905-0134370;
PR	18-MAY-1999;	9905-0134768;
PR	19-MAY-1999;	9905-0134941;
PR	20-MAY-1999;	9905-0135124;
PR	21-MAY-1999;	9905-0135353;
PR	24-MAY-1999;	9905-0135628;
PR	25-MAY-1999;	9905-0136021;
PR	27-MAY-1999;	9905-0136392;
PR	28-MAY-1999;	9905-0136782;
PR	01-JUN-1999;	9905-0137222;
PR	03-JUN-1999;	9905-0137528;
PR	04-JUN-1999;	9905-0137502;
PR	07-JUN-1999;	9905-0137722;
PR	08-JUN-1999;	9905-0138094;
PR	10-JUN-1999;	9905-0138540;
PR	14-JUN-1999;	9905-0139119;
PR	16-JUN-1999;	9905-0139452;
PR	16-JUN-1999;	9905-0139453;
PR	17-JUN-1999;	9905-0139452;
PR	18-JUN-1999;	9905-0139455;
PR	18-JUN-1999;	9905-0139456;
PR	18-JUN-1999;	9905-0139457;
PR	18-JUN-1999;	9905-0139458;
PR	18-JUN-1999;	9905-0139459;
PR	18-JUN-1999;	9905-0139460;
PR	18-JUN-1999;	9905-0139461;
PR	18-JUN-1999;	9905-0139462;
PR	18-JUN-1999;	9905-0139463;
PR	18-JUN-1999;	9905-0139750;
PR	18-JUN-1999;	9905-0139763;
PR	21-JUN-1999;	9905-0139817;
PR	22-JUN-1999;	9905-0139899;
PR	23-JUN-1999;	9905-0140353;
PR	23-JUN-1999;	9905-0140354;
PR	24-JUN-1999;	9905-0140695;
PR	28-JUN-1999;	9905-0140823;
PR	29-JUN-1999;	9905-0140991;
PR	30-JUN-1999;	9905-0141287;
PR	01-JUL-1999;	9905-0141842;
PR	01-JUL-1999;	9905-0142154;
PR	02-JUL-1999;	9905-0142055;
PR	06-JUL-1999;	9905-0142390;
PR	08-JUL-1999;	9905-0142803;
PR	09-JUL-1999;	9905-0142920;
PR	12-JUL-1999;	9905-0142972;
PR	13-JUL-1999;	9905-0143542;
PR	14-JUL-1999;	9905-0143624;
PR	15-JUL-1999;	9905-0144085;
PR	16-JUL-1999;	9905-0144085;
PR	16-JUL-1999;	9905-0144086;
PR	19-JUL-1999;	9905-0144325;
PR	19-JUL-1999;	9905-0144331;
PR	19-JUL-1999;	9905-0144332;
PR	19-JUL-1999;	9905-0144333;
PR	19-JUL-1999;	9905-0144334;
PR	19-JUL-1999;	9905-0144335;
PR	20-JUL-1999;	9905-0144632;
PR	20-JUL-1999;	9905-0144684;
PR	21-JUL-1999;	9905-0144814;
PR	21-JUL-1999;	9905-0145086;
PR	21-JUL-1999;	9905-0145088;
PR	22-JUL-1999;	9905-0145087;
PR	22-JUL-1999;	9905-0145089;
PR	22-JUL-1999;	9905-0145192;
PR	23-JUL-1999;	9905-0145145;
PR	23-JUL-1999;	9905-0145218;
PR	23-JUL-1999;	9905-0145224;
PR	26-JUL-1999;	9905-0145276;
PR	27-JUL-1999;	9905-0145513;
PR	27-JUL-1999;	9905-0145818;
PR	28-JUL-1999;	9905-0145919;
PR	28-JUL-1999;	9905-0145951;
PR	02-AUG-1999;	9905-0146388;
PR	02-AUG-1999;	9905-0146388;
PR	02-AUG-1999;	9905-0146589;
PR	03-AUG-1999;	9905-0147038;
PR	04-AUG-1999;	9905-0147204;
PR	04-AUG-1999;	9905-0147302;
PR	05-AUG-1999;	9905-0147192;
PR	05-AUG-1999;	9905-0147260;
PR	06-AUG-1999;	9905-0147263;
PR	06-AUG-1999;	9905-0147416;
PR	09-AUG-1999;	9905-0147493;
PR	09-AUG-1999;	990

PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145224.
PR 23-JUL-1999; 99US-0145228.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0152363.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 23-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 14-OCT-1999; 99US-0159639.

PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 90.5%; Score 1139; DB 21; Length 241;
Best Local Similarity 89.0%; Pred. No. 5.3e-106;
Matches 202; Conservative 12; Mismatches 13; Indels 0; Gaps 0;

QY 2 GGGWVAHAATFYGGDASGTMGACGCGNLYSGYGTNTAALSTALFNNGLSCGCFEIR 61
Db 15 GGGWVAHAATFYGGDASGTMGACGCGNLYSGYGTNTAALSTALFNNGLSCGCFEIR 74

QY 62 CONDGKMCLEPGSTIVTATNFCPPNNALPNNAGWCNPPQOHFDLSQVFORIAOYRAGIV 121
Db 75 CENDGKMCLEPGSTIVTATNFCPPNNALPNNAGWCNPPLEHFDLQPVQRIAOYRAGIV 134

QY 122 PVAYRRVPCVRRGIRFTINGHSYFNLVLTIVNGAGDVHASAVKSGRTGQWQAMSRNMQ 181
Db 135 PVAYRRVPCVRRGIRFTINGHSYFNLVLTIVNGAGDVHASAIKSGRTGQWQAMSRNMQ 194

QY 182 NMOSNSTYLNQSLSEKVTTSDDGOTIYSNNKXANAGWRFQGTFTGAHR 228
Db 195 NMOSNSTYLNQSLSEKVTTSDDGOTIYSNNKXANAGWRFQGTFTGAHR 241

RESULT 3
ID AAG51633 standard; Protein; 241 AA.
AC AAG51633;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 65551.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
OS
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0123788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.

PT Expansin proteins which alter the mechanical strength of
PT poly:saccharide(s) - useful in paper mfr. and recycling
XX
PS Disclosure: page 30-31; 60pp; English.

XX Expansins are a novel class of proteins that catalyse the extension
CC of plant cell walls and the weakening of the hydrogen bonds in
CC cellulose. 2 Expansins (AAR94528 and AAR94529) have been identified in
CC rice and 3 in Arabidopsis (AAR94530-32). A cDNA clone (AAR13320)
CC coding for cucumber expansin 29 (AAR94527) has been obtd. Expansins
CC can be used e.g. in the mfr., de-linking and recycling of paper, in
CC the textile industry, to aid delignification processes, to alter gel
CC mechanical strength, etc.

XX Sequence 228 AA:

Query Match 99.8%; Score 1257; DB 17; Length 228;
Best Local Similarity 100.0%; Pred. No. 7.2e-118;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AGCGWVNAIATFYGGDASGTGACGYNLYSGYGTNTALSTALFNNGLSGACFEI 60
DB 1 AGCGWVNAIATFYGGDASGTGACGYNLYSGYGTNTALSTALFNNGLSGACFEI 60
OY 61 RCQNDGKWCLEPGSIYVVTATNFCPPNNALPNNAGCNCPPQHFDSLQPVFORIAQYRAGI 120
DB 61 RCQNDGKWCLEPGSIYVVTATNFCPPNNALPNNAGCNCPPQHFDSLQPVFORIAQYRAGI 120
OY 121 VVAVARVPCVRCGRFTINGHSYFNVLVLTNNVGAGGVHSAMKGSRTGQAMSRNG 180
DB 121 VVAVARVPCVRCGRFTINGHSYFNVLVLTNNVGAGGVHSAMKGSRTGQAMSRNG 180
OY 181 QNMOSNLYLNGOSLSFEKVTTSQDQIVSNXNAGMSFCQFTTGAHVR 228
DB 181 QNMOSNLYLNGOSLSFEKVTTSQDQIVSNXNAGMSFCQFTTGAHVR 228

RESULT 2
AAG06546
ID AAG06546 standard; Protein; 241 AA.

XX AAG06546;

XX 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 3357.

XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX Arabidopsis thaliana.

XX EPI033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137529.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140655.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142380.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 11, 2002, 14:49:55 ; Search time 25.2399 Seconds
(without alignments)
1003.367 Million cell updates/sec

Title: US-09-896-301-2

Perfect score: 1259

Sequence: 1 AGGGWNAHAFYGGDASG.....NNXNAGWSPGQTGTGAHVR 228

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: A_Geneseq_032802.*
2: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
3: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
4: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
5: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
6: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
7: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
8: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
9: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
10: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
11: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
12: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
13: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
14: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
15: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
16: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
17: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
18: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
19: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
20: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
21: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
22: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
23: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1257	99.8	228	17	AA894528
2	1139	90.5	241	21	AA606546
3	1139	90.5	241	21	AA631633
4	1139	90.5	241	21	AA651647
5	1139	90.5	249	21	AA606545
6	1139	90.5	249	21	AA651632
7	1139	90.5	249	21	AA651646
8	1139	90.5	259	21	AA606544
9	1139	90.5	259	21	AA651631
10	1139	90.5	280	21	AA651645
11	1115	88.6	251	21	AA633852

12	1115	88.6	251	21	AA643343	Arabidopsis thalia
13	1115	88.6	253	21	AA623851	Arabidopsis thalia
14	1115	88.6	253	21	AA623342	Arabidopsis thalia
15	1115	88.6	281	21	AA643341	Arabidopsis thalia
16	1115	88.6	282	21	AA623850	Arabidopsis thalia
17	1046	83.1	250	22	AA600414	Tomato seed expans
18	1026	81.5	227	17	AA694527	Cucumber expansin-
19	990	78.6	241	21	AA605452	Arabidopsis thalia
20	990	78.6	249	21	AA605451	Arabidopsis thalia
21	990	78.2	255	21	AA636570	Arabidopsis thalia
22	985	78.2	242	21	AA636569	Arabidopsis thalia
23	985	78.2	255	21	AA636568	Arabidopsis thalia
24	976.5	77.6	253	21	AA625443	Arabidopsis thalia
25	976.5	77.6	253	21	AA646483	Arabidopsis thalia
26	976.5	77.6	253	21	AA694532	Arabidopsis thalia
27	959.5	76.2	227	17	AA694532	Arabidopsis thalia
28	955	75.9	257	21	AA636445	Arabidopsis thalia
29	930.5	73.9	221	21	AA615694	Arabidopsis thalia
30	930.5	73.9	221	21	AA651013	Arabidopsis thalia
31	930.5	73.9	255	21	AA615693	Arabidopsis thalia
32	930.5	73.9	255	21	AA651012	Arabidopsis thalia
33	921	73.2	262	21	AA629831	Arabidopsis thalia
34	921	73.2	273	21	AA629930	Arabidopsis thalia
35	916	72.8	253	22	AA600412	Tomato seed expans
36	913	72.5	250	21	AA609622	Arabidopsis thalia
37	913	72.5	258	21	AA609621	Arabidopsis thalia
38	913	72.5	280	21	AA609620	Arabidopsis thalia
39	908	72.1	255	21	AA630325	Arabidopsis thalia
40	908	72.1	257	21	AA630324	Arabidopsis thalia
41	898.5	71.4	222	17	AA694529	Rice expansin. Or
42	890.5	70.7	207	21	AA625444	Arabidopsis thalia
43	890.5	70.7	207	21	AA646484	Arabidopsis thalia
44	859	68.2	225	17	AA694531	Arabidopsis expans
45	856	68.0	210	21	AA636446	Arabidopsis thalia

ALIGNMENTS

RESULT 1	AA894528	standard; Protein; 228 AA.
XX	AA894528;	
AC	AA894528;	
XX		
DT	08-JUL-1996	(first entry)
XX		
DE	Rice expansin.	
XX		
KW	Expansin; plant cell wall; cellulose; paper recycling; de-inking;	
KM	polysaccharide; rice.	
XX		
OS	Oryza sativa.	
XX		
FH	Key	Location/Qualifiers
FT	Misc-difference 211	
FT		/note- "unidentified amino acid"
XX		
PN	AAU540262-A.	
XX		
PD	04-APR-1996.	
XX		
PE	12-MAY-1994;	94AU-0068320.
XX		
PR	12-MAY-1995;	95US-0440517.
XX		
PR	12-MAY-1993;	93US-0060944.
XX		
PA	(PENN-) PENN STATE RES FOUND.	
XX		
PI	Cosgrove DJ, McQueen-Mason S;	
XX		
DR	WPI; 1996-201150/21.	
XX		


```

Query Match      58.3%; Score 733.5; DB 4; Length 227;
Best Local Similarity 63.8%; Pred. No. 3.1e-66;
Matches 136; Conservative 22; Mismatches 52; Indels 3; Gaps 2;

QY 7 NAHATPEYGGDASGTMGAGCGYNLYSOGYGTNTALSTALFNNGLSGACFEIRCO-ND 65
   :| ||||| | | ||||| ||||| : ||| : ||||| : ||| : | :
Db 7 SAFATPEYGGKDGSCGTMGAGCGYNLYNAGYGLYNALSSALFNDGAMCGACTTTCDSQ 66
   :| ||||| | | ||||| ||||| : ||| : ||||| : ||| : | :
QY 66 GKACLPG--SIVVTATNFCEPPNNALPNNAGGWCNPPQOHFDLSQPVFQRIAYRAGIVPV 123
   ||| || | | : ||||| | | ||||| ||||| : ||| : ||||| : ||| : | :
Db 67 TKMCKPQGNMSTTTATNLCKPNNALPNSGCMCPPLXHFDMSPAMENIAYVQAGIVPV 126
   ||| || | | : ||||| | | ||||| ||||| : ||| : ||||| : ||| : | :
QY 124 AYRRVPCVRRGGIRFTINGHSYFNVLITVVGAGDVHSAMVKGSRTGQMSRNNGQNW 183
   ||||| | | ||||| : ||| | | : ||||| : | : ||| ||| ||||| ||
Db 127 NTKRVPQKRSGLRFAISGHDYELVTVTVGGSGVVAQMSIKGSNTGWMMSRNNGANN 186
   ||| : || ||||| | | : : : ||| |
QY 184 QSNSYLNQSLSPFKVTTSDGQTIIVSNXXANAGW 216
   ||| : || ||||| | | : : : ||| |
Db 187 QSNAYLAGSLSFIVQLDQGRKVTANNXXAPXNW 219
   ||| : || ||||| | | : : : ||| |

```

Search completed: October 11, 2002, 15:02:04
 Job time : 8.91808 secs

Query Match 61.3%: Score 772; DB 2; Length 179;
Best Local Similarity 77.4%: Pred. No. 3.1e-70;
Matches 137; Conservative 14; Mismatches 24; Indels 2; Gaps 1;

QY 10 ATFFYGGDASGTMGACGYNLVSGYGTNTALSTALFNNGLSGACFEIRCONDGKMC 69
DB 3 ATFFYGGDASGTMGACGYNLVSGYGTNTALSTALFNNGLSGACFEIRCONDGKMC 62
QY 70 LFGS--IIVTATNFCPPNNALPNNAGMCNPPQHFDSLQPVFORIAOYRAGIYPAVYR 127
DB 63 HPGSPCIFTTATNFCPPNNALPNDNGMCNLPRTHEFLAMPFLKIAEYRAGIYPAVYR 122
QY 128 VPCVRGCGIRFTINGSHYFNVLITNVGAGDVHSAMVKGSRGTGQAMSRMGNQW 184
DB 123 VPCVRGCGIRFTINGSHYFNVLITNVGAGDVHSAMVKGSRGTGQAMSRMGNQW 179

RESULT 13
US-09-362-642-6
Sequence 6, Application US/09362642
Patent No. 6350935

GENERAL INFORMATION:
APPLICANT: Bennett, Alan B.
APPLICANT: Rose, Jocelyn K.C.
TITLE OF INVENTION: The Regents of the University of California
TITLE OF INVENTION: Fruit-Specific and Ripening Regulation Expansin Genes
FILE REFERENCE: 023070-078210US
CURRENT APPLICATION NUMBER: US/09/362, 642
CURRENT FILING DATE: 1999-07-27
NUMBER OF SEQ ID NOS: 8
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 6
LENGTH: 179
TYPE: PRT
ORGANISM: Cucumis melo
US-09-362-642-6

Query Match 61.3%: Score 772; DB 4; Length 179;
Best Local Similarity 77.4%: Pred. No. 3.1e-70;
Matches 137; Conservative 14; Mismatches 24; Indels 2; Gaps 1;

QY 10 ATFFYGGDASGTMGACGYNLVSGYGTNTALSTALFNNGLSGACFEIRCONDGKMC 69
DB 3 ATFFYGGDASGTMGACGYNLVSGYGTNTALSTALFNNGLSGACFEIRCONDGKMC 62
QY 70 LFGS--IIVTATNFCPPNNALPNNAGMCNPPQHFDSLQPVFORIAOYRAGIYPAVYR 127
DB 63 HPGSPCIFTTATNFCPPNNALPNDNGMCNLPRTHEFLAMPFLKIAEYRAGIYPAVYR 122
QY 128 VPCVRGCGIRFTINGSHYFNVLITNVGAGDVHSAMVKGSRGTGQAMSRMGNQW 184
DB 123 VPCVRGCGIRFTINGSHYFNVLITNVGAGDVHSAMVKGSRGTGQAMSRMGNQW 179

RESULT 14
US-08-440-517A-4
Sequence 4, Application US/08440517A
Patent No. 5959082

GENERAL INFORMATION:
APPLICANT: COSGROVE, DANIEL J.;
APPLICANT: GUILTINAN, MARK;
APPLICANT: SHERBAN, TATYANA;
TITLE OF INVENTION: PURIFIED EXPANSIN PROTEINS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: INTELLECTUAL PROPERTY OFFICE, THE
STREET: 113 TECHNOLOGY CENTER
CITY: UNIVERSITY PARK
STATE: PENNSYLVANIA

COUNTRY: UNITED STATES OF AMERICA
ZIP: 16802-7000
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: NEC 286
OPERATING SYSTEM: DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/440, 517A
FILING DATE:

CLASSIFICATION: 530
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 227
TYPE: AMINO ACID
TOPOLOGY: UNKNOWN

US-08-440-517A-4

Query Match 58.3%: Score 733.5; DB 2; Length 227;
Best Local Similarity 63.8%: Pred. No. 3.1e-66;
Matches 136; Conservative 22; Mismatches 52; Indels 3; Gaps 2;

QY 7 NAHATFYGDDASGTMGACGYNLVSGYGTNTALSTALFNNGLSGACFEIRCO-ND 65
DB 7 SAFAIFYGKDSCTMGACGYNLVAGGLYNALSLFNDGAMCGCYTTCDT50 66
QY 66 GKWLPG--IIVTATNFCPPNNALPNNAGMCNPPQHFDSLQPVFORIAOYRAGIY 123
DB 67 TKWCPGNSITITATNLKAPNNALPNSGKCNPLXHPDMOPAMENIAVYQAGIY 126
QY 124 AYRRVPCVRGCGIRFTINGSHYFNVLITNVGAGDVHSAMVKGSRGTGQAMSRMGNQW 183
DB 127 NYKRVFXQSGGIRFALISGHDFELVTVNVGSGVVAQMSIGKSTGMAMSRMGNQW 186
QY 184 QSNXYLNGQSLSPKVTTSDDQITVSNXANAGW 216
DB 187 QSNAYLAGQSLSPVQLDGRKVTAMNXPXNW 219

RESULT 15
US-09-092-160-4
Sequence 4, Application US/09092160C
Patent No. 6235466

GENERAL INFORMATION:
APPLICANT: Cosgrove, Daniel J
APPLICANT: McQueen-Mason, Simon
APPLICANT: Gullitnan, Mark J
APPLICANT: Sherban, Tatyana
TITLE OF INVENTION: PURIFIED EXPANSIN PROTEINS
FILE REFERENCE: 1194/1C114US3
CURRENT APPLICATION NUMBER: US/09/092, 160C
CURRENT FILING DATE: 1998-06-05
EARLIER APPLICATION NUMBER: 08/440, 517
EARLIER FILING DATE: 1995-05-12
EARLIER APPLICATION NUMBER: 08/242, 090
EARLIER FILING DATE: 1994-05-12
EARLIER APPLICATION NUMBER: 08/060, 944
NUMBER OF SEQ ID NOS: 7
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 4
LENGTH: 227
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Arabidopsis
NAME/KEY: UNSURE
LOCATION: (2)-(227)
OTHER INFORMATION: Xaa is unknown or other.
US-09-092-160-4

US-08-440-517A-6

Query Match	76.2%;	Score 959.5;	DB 2;	Length 226;
Best Local Similarity	75.0%;	Pred. No. 5.6e-89;		
Matches 165; Conservative	23;	Mismatches 31;	Indels 1;	Gaps 1.

QY	5	WVNNAHATFYGGGGDASGTMGACGCGNGLYSOSYGTNTAALSTALENNLSGACGEIRCON	64
	1		
	5	WOSHATFTYGGGGDASGTMGTGCTGIGNLSTCY-TNTAALSTLVLENDACRSCTELCDN	63
Db			
QY	65	DGKMCLPGSIYVLTATNTFCPPNNALPNNAGGWCNCPQOHFDLSQEFORIAQYRAGIYPA	124
	1		
	64	DGQMCPLPSYVLTATNLCPRYALPNDDDGGWCNRPREDHMAEPALQIGYRAGIYPS	123
Db			
QY	125	YRRPDCVARRGIRFTINGHSTFENLVLTNNVGAGDVHSAMTKSGRTTQMSRRMGQNMQ	184
	1		
	124	YRRPDCVAKKGIRFTINGHSTFENLVLTNNAAGPDVDVSYSIKGSSGTGMPMSRRMGQNMQ	183
Db			
QY	185	SNSLYNGGSLSEKVTTSQGTIVSNKNYANGMSGQDFTG	224
	1		
Db	184	SNSLYDGSSLSFYQAVSDGRVTNNVYPMQDYGQTFEG	223

RESULT 5
US-09-092-160-6
; Sequence 6, Application US/09092160C

```

GENERAL INFORMATION:
APPLICANT: Cosgrove, Daniel J
APPLICANT: McQueen-Mason, Simon
APPLICANT: Gullinan, Mark J
APPLICANT: Shcherban, Tatyana
APPLICANT: Shi, Jun
TITLE OF INVENTION: PURIFIED EXPANSIN PROTEINS
FILE REFERENCE: 1194/1C114US3
CURRENT APPLICATION NUMBER: US/09/092,160C
CURRENT FILING DATE: 1998-06-05
EARLIER APPLICATION NUMBER: 08/440,517
EARLIER FILING DATE: 1995-05-12
EARLIER APPLICATION NUMBER: 08/742,090
EARLIER FILING DATE: 1994-05-12
EARLIER APPLICATION NUMBER: 08/060,944
EARLIER FILING DATE: 1993-05-12
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 226
TYPE: PR1
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Arabidopsis
US-09-092-160-6

```

Query Match	76.2%;	Score 959.5;	DB 4;	Length 226;
Best Local Similarity	75.0%;	Pred. No. 5.6e-89;		
Matches 165;	Conservative 23;	Mismatches 31;	Indels 1;	Gaps 1

[illegible]

RESULT 6
US-08-440-517A-3
; Sequence 3, Application US/08440517A
; Patent No. 5959082

1 GENERAL INFORMATION:
 2 APPLICANT: COSGROVE, DANIEL J.;
 3 APPLICANT: SUGHTJINAN, MARK;
 4 APPLICANT: SUGHERBAN, TATYANA;
 5 APPLICANT: SHI, JUN
 6 TITLE OF INVENTION: PURIFIED EXPANSIN PROTEINS
 7 NUMBER OF SEQUENCES: 6
 8 CORRESPONDENCE ADDRESS:
 9 ADDRESSEE: INTELLECTUAL PROPERTY OFFICE, THE
 10 ADDRESSEE: PENNSYLVANIA STATE UNIVERSITY
 11 STREET: 113 TECHNOLOGY CENTER
 12 CITY: UNIVERSITY PARK
 13 STATE: PENNSYLVANIA
 14 COUNTRY: UNITED STATES OF AMERICA
 15 ZIP: 16802-7000
 16
 17 COMPUTER READABLE FORM:
 18 MEDIUM TYPE: FLOPPY DISK
 19 COMPUTER: NEC 286
 20 OPERATING SYSTEM: DOS
 21 SOFTWARE: WORDPERECT 5.1
 22 CURRENT APPLICATION DATA:
 23 APPLICATION NUMBER: US/08/440,517A
 24 FILING DATE:
 25 CLASSIFICATION: 530
 26 INFORMATION FOR SEQ ID NO: 3:
 27 SEQUENCE CHARACTERISTICS:
 28 LENGTH: 222
 29 TYPE: AMINO ACID
 30 TOPOLOGY: UNKNOWN
 31 US-08-440-517A-3

Query Match	71.48;	Score 898.5;	DB 2;	Length 222;
Best Local Similarity	72.28;	Pred. No. 7.6e-83;		
Matches 164;	Conservative 19;	Mismatches 35;	Indels 9;	Gaps 2

QY	3	GGWVNAHAFYCGDASGMSGACGYGVTSQGGTNTALSTALFENNNLSGCAFEELRC	62
Db	3	GPMVNAHAFYKXGDAXXTMGACGYGNLSQGTLETTALSTALFPDGLSGCAXEFLMC	62
QY	63	QNGKMCWCEG-SIVYATATNECPNPNNALPNNAGWCNPPQOHEPDSQPVFORIAOYRACIV	121
Db	63	VNDPQMCINGRSIIVYATATNECP-----GGACDPNPNNHPLDSQPIYEALYKSGII	114
QY	122	PVATRRVPCVRRGGTREFITINGHSYFNLVLTNNGGAGDVHSAMVKSRTGWAQMSRMNQ	181
Db	115	PVMRRVRCKRSGGTRFTINGHSYFNLVLTNNGGAGDVHSAMKSRTKWQMSRMNQ	174
QY	182	NMQSNSYLNGQSLSEKVTYTSQGITVSNMKNANMGWSGCFQFCAHNR	228
Db	175	NMQSNSYLNGQSLSEKVTYTSQGRSVSPNAPPPWSGQJYGGQFR	221

```

: RESUME 2
: US-09-092-160-3
: Sequence 3, Application US/09092160C
: Patent No. 6255466
: GENERAL INFORMATION:
: APPLICANT: Cosgrove, Daniel J
: APPLICANT: McQueen-Wason, Simon
: APPLICANT: Gullinan, Mark J
: APPLICANT: Shecharban, Talyana
: APPLICANT: Shi, Jun
: TITLE OF INVENTION: PURIFIED EXPANSIN PROTEINS
: FILE REFERENCE: 1194/IC114US3
: CURRENT APPLICATION NUMBER: US/09/092,160C
: CURRENT FILING DATE: 1998-06-05
: EARLIER APPLICATION NUMBER: 08/440,517

```

QY 181 QMOSNLYNGQSLSFKVTTSDDQTISSNNXNAGWMSFGQFTGAHVR 228
DB 181 QMOSNLYNGQSLSFKVTTSDDQTISSNNXNAGWMSFGQFTGAHVR 228

RESULT 2

US-09-092-160-2

Sequence 2, Application US/09092160C

Patent No. 6255466

GENERAL INFORMATION:

APPLICANT: Cosgrove, Daniel J

APPLICANT: McQueen-Mason, Simon

APPLICANT: Gullitman, Mark J

APPLICANT: Shcherban, Tatyana

TITLE OF INVENTION: PURIFIED EXPANSIN PROTEINS

FILE REFERENCE: 1194/IC114US3

CURRENT APPLICATION NUMBER: US/09/092,160C

EARLIER FILING DATE: 1998-06-05

EARLIER APPLICATION NUMBER: 08/440,517

EARLIER FILING DATE: 1995-05-12

EARLIER APPLICATION NUMBER: 08/242,090

EARLIER FILING DATE: 1994-05-12

EARLIER APPLICATION NUMBER: 08/060,944

EARLIER FILING DATE: 1993-05-12

NUMBER OF SEQ ID NOS: 7

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 2

LENGTH: 228

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: rice expansin

NAME/KEY: UNSURE

LOCATION: 211

OTHER INFORMATION: Xaa is unknown or other.

US-09-092-160-2

Query Match

Best Local Similarity 99.8%; Score 1257; DB 4; Length 228;

Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGCVNAHATFYGGDASGTMGACGYGNYLSQGYGTNTAALSTALFNNGLSGACFEI 60
DB 1 AGGCVNAHATFYGGDASGTMGACGYGNYLSQGYGTNTAALSTALFNNGLSGACFEI 60
QY 61 RQNDGKCLGSGIYVATNCPFPNNALPNNAGWCNPPQHFDSQVFORIAYRAGI 120
DB 61 RQNDGKCLGSGIYVATNCPFPNNALPNNAGWCNPPQHFDSQVFORIAYRAGI 120
QY 121 VPVAVRVCVRGIRFTINGHSYFNVLITVNGAGADVSAWVKSRTGQMSRNMWG 180
DB 121 VPVAVRVCVRGIRFTINGHSYFNVLITVNGAGADVSAWVKSRTGQMSRNMWG 180
QY 181 QMOSNLYNGQSLSFKVTTSDDQTISSNNXNAGWMSFGQFTGAHVR 228
DB 181 QMOSNLYNGQSLSFKVTTSDDQTISSNNXNAGWMSFGQFTGAHVR 228

RESULT 3

US-09-092-160-7

Sequence 7, Application US/09092160C

Patent No. 6255466

GENERAL INFORMATION:

APPLICANT: Cosgrove, Daniel J

APPLICANT: McQueen-Mason, Simon

APPLICANT: Gullitman, Mark J

APPLICANT: Shcherban, Tatyana

TITLE OF INVENTION: PURIFIED EXPANSIN PROTEINS

FILE REFERENCE: 1194/IC114US3

CURRENT APPLICATION NUMBER: US/09/092,160C

EARLIER FILING DATE: 1998-06-05

EARLIER APPLICATION NUMBER: 08/440,517

EARLIER FILING DATE: 1995-05-12

EARLIER APPLICATION NUMBER: 08/242,090

EARLIER FILING DATE: 1994-05-12

EARLIER APPLICATION NUMBER: 08/060,944

EARLIER FILING DATE: 1993-05-12

NUMBER OF SEQ ID NOS: 7

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 7

LENGTH: 227

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: cucumber

NAME/KEY: UNSURE

LOCATION: 211

OTHER INFORMATION: Xaa is unknown or other.

US-09-092-160-7

QY 3 GGVNAHATFYGGDASGTMGACGYGNYLSQGYGTNTAALSTALFNNGLSGACFEI 62
DB 3 GGVNAHATFYGGDASGTMGACGYGNYLSQGYGTNTAALSTALFNNGLSGACFEI 62
QY 63 QNDGKCLGSGIYVATNCPFPNNALPNNAGWCNPPQHFDSQVFORIAYRAGI 122
DB 63 QNDGKCLGSGIYVATNCPFPNNALPNNAGWCNPPQHFDSQVFORIAYRAGI 122
QY 123 VPVAVRVCVRGIRFTINGHSYFNVLITVNGAGADVSAWVKSRTGQMSRNMWG 182
DB 123 VPVAVRVCVRGIRFTINGHSYFNVLITVNGAGADVSAWVKSRTGQMSRNMWG 182
QY 183 QMOSNLYNGQSLSFKVTTSDDQTISSNNXNAGWMSFGQFTGAHVR 224
DB 183 QMOSNLYNGQSLSFKVTTSDDQTISSNNXNAGWMSFGQFTGAHVR 224

Query Match

Best Local Similarity 81.1%; Score 1021; DB 4; Length 227;

Matches 175; Conservative 23; Mismatches 24; Indels 0; Gaps 0;

US-08-440-517A-6
Sequence 6, Application US/08440517A
Patent No. 5959082
GENERAL INFORMATION:
APPLICANT: COSGROVE, DANIEL J.
APPLICANT: GULLITMAN, MARK J.
APPLICANT: SCHERBAN, TATYANA;
TITLE OF INVENTION: PURIFIED EXPANSIN PROTEINS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: INTELLECTUAL PROPERTY OFFICE, THE
ADDRESS: PENNSYLVANIA STATE UNIVERSITY
STREET: 113 TECHNOLOGY CENTER
CITY: UNIVERSITY PARK
STATE: PENNSYLVANIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 16802-7000
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: NEC 286
OPERATING SYSTEM: DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/440,517A
FILING DATE:
CLASSIFICATION: 530
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 226
TYPE: AMINO ACID
TOPOLOGY: UNKNOWN

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 11, 2002, 14:54:05 ; Search time 8.91808 Seconds
(without alignments)
624.466 Million cell updates/sec

Title: US-09-896-301-2

Perfect score: 1259
Sequence: 1 AGCGWVNAHATFGGDSAG.....NNXANAGNSFGQTFGAHVR 228

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Issued Patents AA:*
1: /cgn2_6/prodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/prodata/2/1aa/PCITUS.COMB.pep:*
6: /cgn2_6/prodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1257	99.8	228	2	US-08-440-517A-2
2	1257	99.8	228	4	US-09-092-160-2
3	1021	81.1	227	4	US-09-092-160-7
4	959.5	76.2	226	2	US-08-440-517A-6
5	959.5	76.2	226	4	US-09-092-160-6
6	898.5	71.4	222	2	US-08-440-517A-3
7	898.5	71.4	222	4	US-09-092-160-3
8	859	68.2	225	2	US-08-440-517A-5
9	859	68.2	225	4	US-09-092-160-5
10	841	66.8	225	4	US-09-362-642-2
11	836	66.4	225	2	US-08-845-539-2
12	772	61.3	179	2	US-08-845-539-6
13	772	61.3	179	4	US-09-362-642-6
14	733.5	58.3	227	2	US-08-440-517A-4
15	733.5	58.3	227	4	US-09-092-160-4
16	711	56.5	167	2	US-08-845-539-4
17	711	56.5	167	4	US-09-362-642-4
18	731	18.3	261	1	US-07-971-096-2
19	231	18.3	261	1	US-08-175-096-2
20	222.5	17.5	246	4	US-08-441-507-21
21	220.5	17.5	272	4	US-08-441-507-15
22	212	16.8	263	1	US-07-971-096-4
23	212	16.8	263	1	US-08-175-096-4
24	212	16.8	263	4	US-08-413-974-6
25	212	16.8	263	4	US-08-434-418-6
26	212	16.8	263	4	US-08-433-288-6
27	212	16.8	263	4	US-08-174-739A-6

28	191.5	15.2	245	4	US-08-441-507-24	Sequence 24, Appl
29	154	12.2	197	4	US-08-441-507-5	Sequence 5, Appl
30	154	12.2	200	4	US-08-441-507-4	Sequence 4, Appl
31	152	12.1	122	4	US-08-441-507-23	Sequence 23, Appl
32	101	8.0	54	3	US-08-750-419A-12	Sequence 12, Appl
33	97.5	7.7	145	4	US-08-413-974-4	Sequence 4, Appl
34	97.5	7.7	145	4	US-08-434-418-4	Sequence 4, Appl
35	97.5	7.7	145	4	US-08-433-288-4	Sequence 4, Appl
36	97.5	7.7	145	4	US-08-174-739A-4	Sequence 4, Appl
37	83	6.6	1477	1	US-08-038-682-4	Sequence 4, Appl
38	83	6.6	1477	1	US-08-302-832-4	Sequence 4, Appl
39	83	6.6	1477	2	US-08-530-198-4	Sequence 4, Appl
40	83	6.6	1477	2	US-08-469-880-4	Sequence 4, Appl
41	83	6.6	1477	2	US-08-728-470-4	Sequence 4, Appl
42	83	6.6	1477	2	US-08-617-697-4	Sequence 4, Appl
43	83	6.6	1477	4	US-08-719-641-4	Sequence 4, Appl
44	82	6.5	1160	3	US-08-808-599A-24	Sequence 24, Appl
45	81	6.4	890	2	US-08-483-101-14	Sequence 14, Appl

ALIGNMENTS

```
RESULT 1
US-08-440-517A-2
; Sequence 2, Application US/08440517A
; Patent No. 5959082
; GENERAL INFORMATION:
; APPLICANT: COSGROVE, DANIEL J.;
; APPLICANT: GULLITMAN, MARK;
; APPLICANT: SCHERBAN, TATYANA;
; APPLICANT: SHI, JUN
; TITLE OF INVENTION: PURIFIED EXPANSIN PROTEINS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INTELLECTUAL PROPERTY OFFICE, THE
; ADDRESSEE: PENNSYLVANIA STATE UNIVERSITY
; STREET: 113 TECHNOLOGY CENTER
; CITY: UNIVERSITY PARK
; STATE: PENNSYLVANIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 16802-7000
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: NEC 286
; OPERATING SYSTEM: DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/440,517A
; FILING DATE:
; CLASSIFICATION: 530
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 228
; TYPE: AMINO ACID
; TOPOLOGY: UNKNOWN
US-08-440-517A-2
Query Match 99.8%; Score 1257; DB 2; Length 228;
Best local similarity 100.0%; Pred. No. 6.5e-119;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 AGCGWVNAHATFGGDSAGTGMGACGYGMLYSOGYGTNTAALSTALFNNGLSCGACFEL 60
1 RCONDKGWCPCPSIVVATNFCPPNNALPNNAGGWCNPPQOHFDSLQPVFORIAQYRAGI 120
1 AGCGWVNAHATFGGDSAGTGMGACGYGMLYSOGYGTNTAALSTALFNNGLSCGACFEL 60
1 RCONDKGWCPCPSIVVATNFCPPNNALPNNAGGWCNPPQOHFDSLQPVFORIAQYRAGI 120
121 VPAVYRRVPCVRRGRIFFTTNGHSYFNLVLTIVNGAGDVHSAVKGSRGTGWQAMSRNMG 180
121 VPAVYRRVPCVRRGRIFFTTNGHSYFNLVLTIVNGAGDVHSAVKGSRGTGWQAMSRNMG 180
```

6

QY 119 GIVPVAARYPCVRRGIRFTINGHSYFNLYLTNVGAGDVHSAWKGSRTOQAMSRN 178
 Db 147 GIVPVSTRVRRPCKRGIRFTINGHRENFNLVLTNVAGADIVRASVKSRTGMSLSRN 206
 QY 179 WGNQMSNSYLNGQSLSFKVTTSDDGOTIVSNXNAGWSFGOTFGAHVR 228
 Db 207 WGNQMSNAYLVGQSLSFRTVTSDDRTSTSMNVGQGFVGVKNFR 256

RESULT 13

T50658
 C:Species: Lycopersicon esculentum (tomato)
 C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 28-Jul-2000
 C:Accession: T50658
 R:Caderas, D.M.
 submitted to the EMBL Data Library, June 1999
 A:Reference number: 225160
 A:Accession: T50658
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-257 <CAD>
 A:Cross-references: EMBL:AJ243340; PIDN:CAB46492.1
 A:Experimental source: cultivar Money Maker
 C:Genetics:
 A:Gene: exp9
 C:Superfamily: expansin

Query Match 74.3%; Score 936; DB 2; Length 257;
 Best Local Similarity 72.2%; Pred. No. 7.8e-74;
 Matches 166; Conservative 23; Mismatches 39; Indels 2; Gaps 1;

QY 1 AGGWNVAHATFYGGDASGTMGACGYNLVSGYGTNTALSTALFNNGLSGACPEI 60
 Db 27 SGGSMQNAHATFYGGDASGTMGACGYNLVSGYGTNTALSTALFNNGLSGACPEI 86
 QY 61 RCNDGKWCPLG--SIVVTATNFCPPNNALPNNAGMCPPOOHFLDLPVQRIAOYRA 118
 Db 87 KCONQPMCHPSPSTLTATNFCPPNNALPNNAGMCPNPTHEFLDLPVQRIAOYRA 146
 QY 119 GIVPVAARYPCVRRGIRFTINGHSYFNLYLTNVGAGDVHSAWKGSRTOQAMSRN 178
 Db 147 GIVPVSTRVRRPCKRGIRFTINGHSYFNLYLTNVGAGADIVRASVKSRTGMSLSRN 206
 QY 179 WGNQMSNSYLNGQSLSFKVTTSDDGOTIVSNXNAGWSFGOTFGAHVR 228
 Db 207 WGNQMSNAYLVGQSLSFRTVTSDDRTSTSMNVGQGFVGVKNFR 256

RESULT 14

T50655
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 28-Jul-2000
 C:Accession: T50655
 R:Shcherban, T.Y.; Shi, J.; Durachko, D.M.; Gultinan, M.J.; McQueen-Mason, S.J.; Shieh, Proc. Natl. Acad. Sci. U.S.A. 92, 9245-9249, 1995
 A:Title: Molecular cloning and sequence analysis of expansins--a highly conserved, multi
 A:Reference number: 214894; MUID:96016146
 A:Accession: T50655
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-255 <SHC>
 A:Cross-references: EMBL:U30478; PIDN:AAB38071.1
 A:Experimental source: variety Columbia
 C:Genetics:
 A:Gene: exp5
 C:Function:
 A:Description: induces extension (creep) in plant cell walls
 C:Superfamily: expansin
 C:Keywords: cell wall

Query Match

73.9%; Score 930.5; DB 2; Length 255;

Best Local Similarity 74.4%; Pred. No. 2.3e-73;
 Matches 169; Conservative 19; Mismatches 30; Indels 9; Gaps 2;
 QY 3 GGVNAHATFYGGDASGTMGACGYNLVSGYGTNTALSTALFNNGLSGACPEI 62
 Db 36 GGVNAHATFYGGDASGTMGACGYNLVSGYGTNTALSTALFNNGLSGACPEI 95
 QY 63 QNDGKWCPLG--SIVVTATNFCPPNNALPNNAGMCPPOOHFLDLPVQRIAOYRA 121
 Db 96 VNDPQMCIKGRSIVVTATNFCPP-----GGACDPPNHHFDLSQPIYEKIALKSGII 147
 QY 122 PVAYRRVPCVRRGIRFTINGHSYFNLYLTNVGAGDVHSAWKGSRTOQAMSRN 181
 Db 148 PVYRRVRRPCKRGIRFTINGHSYFNLYLTNVGAGADIVRASVKSRTGMSLSRN 207
 QY 182 NMQMSNSYLNGQSLSFKVTTSDDGOTIVSNXNAGWSFGOTFGAHVR 228
 Db 208 NMQMSNSYLNGQSLSFRTVTSDDRTSTSMNVGQGFVGVKNFR 254

RESULT 15

T47689
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 26-May-2000
 C:Accession: T47689
 R:Benes, V.; Wurmbach, E.; Drzonek, H.; Ansoerge, W.; Mewes, H.W.; Lemcke, K.; Mayer, submitted to the Protein Sequence Database, February 2000
 A:Reference number: 224472
 A:Accession: T47689
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-260 <BEN>
 A:Cross-references: EMBL:ALJ32975
 A:Experimental source: cultivar Columbia; BAC clone T22E16
 C:Genetics:
 A:Map position: 3
 A:introns: 52/1; 158/2
 A:Note: T22E16.160
 C:Superfamily: expansin

Query Match 73.8%; Score 929; DB 2; Length 260;
 Best Local Similarity 71.7%; Pred. No. 3.2e-73;
 Matches 163; Conservative 25; Mismatches 38; Indels 2; Gaps 1;

QY 1 AGGWNVAHATFYGGDASGTMGACGYNLVSGYGTNTALSTALFNNGLSGACPEI 60
 Db 30 SGGSMQNAHATFYGGDASGTMGACGYNLVSGYGTNTALSTALFNNGLSGACPEI 89
 QY 61 RCNDGKWCPLG--SIVVTATNFCPPNNALPNNAGMCPPOOHFLDLPVQRIAOYRA 118
 Db 90 KCVNPKMCHPSPSTLTATNFCPPNNALPNNAGMCPNPTHEFLDLPVQRIAOYRA 149
 QY 119 GIVPVAARYPCVRRGIRFTINGHSYFNLYLTNVGAGDVHSAWKGSRTOQAMSRN 178
 Db 150 GIVPVSTRVRRPCKRGIRFTINGHRENFNLVLTNVAGADIVRASVKSRTGMSLSRN 209
 QY 179 WGNQMSNSYLNGQSLSFKVTTSDDGOTIVSNXNAGWSFGOTFGAHVR 228
 Db 210 WGNQMSNAYLVGQSLSFRTVTSDDRTSTSMNVGQGFVGVKNFR 259

Search completed: October 11, 2002, 15:00:59
 Job time : 12.4421 secs

Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.

A:Reference number: AB4420; MUID:20083487

A:Accession: C84444

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-248 <STO>

A:Cross-references: GB:AE002093; NID:g3461833; PIDN:ACJ2927.1; GSPDB:GN00139

C:Genetics:

A:Gene: AC2903090

A:Map position: 2

C:Superfamily: expansin

Query Match 88.6%; Score 1115; DB 2; Length 248;

Best Local Similarity 87.1%; Pred. No. 2,3e-89;

Matches 196; Conservative 13; Mismatches 16; Indels 0; Gaps 0;

4 GGNVNAHTFYGGDASGTMGACGYGNTLSQGYGTNTAALSTALFNNGISCGACFEIRCO 63

24 GGNVNAHTFYGGDASGTMGACGYGNTLSQGYGTNTAALSTALFNNGISCGACFEIRCO 83

64 NDGKWCPLPGSIVYATNFPCPNNALPNNAGWCNPPQOHFDLSQPYFORIAQYRAGIVP 123

84 SDGAMCLPGALIVYATNFPCPNNALPNNAGWCNPPRLHFDLSQPYFORIAQYRAGIVP 143

124 AYRRVPCVRKRGIRFTINGHSYFNLVLTITNVGAGDVHSAMVKGSRGTGQAMSRNMGON 183

144 SYRRVPCMRKRGIRFTINGHSYFNLVLTITNVGAGDVHSAMVKGSRGTGQAMSRNMGON 203

184 QSNSTYLNQGSLSFKYVTSDDGTTVSNXNANAGSFGOTFTGCAHV 228

204 QSNSTYLNQGSLSFKYVTSDDGTTVSNXNANAGSFGOTFTGCAHV 248

RESULT 3

109821

expansin (clone pteexp3) - loblolly pine (fragment)

C:Species: Pinus taeda (loblolly pine)

C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 26-May-2000

C:Accession: T09821

R:Hutchinson, K.W.; Singer, P.B.; Diaz-Sala, C.; Greenwood, M.S.

A:Description: Expansins are conserved in conifers and expressed in response to exogenous

A:Reference number: Z16866

A:Accession: T09821

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-232 <HUT>

A:Cross-references: EMBL:U64891; NID:g1778100; PID:g1778101

A:Experimental source: clone pteexp3

C:Superfamily: expansin

Query Match 83.2%; Score 1047; DB 2; Length 232;

Best Local Similarity 79.1%; Pred. No. 1.7e-83;

Matches 178; Conservative 26; Mismatches 21; Indels 0; Gaps 0;

3 GGNVNAHTFYGGDASGTMGACGYGNTLSQGYGTNTAALSTALFNNGISCGACFEIRCO 62

8 GGNVNAHTFYGGDASGTMGACGYGNTLSQGYGTNTAALSTALFNNGISCGACFEIRCO 67

63 QNDGKWCPLPGSIVYATNFPCPNNALPNNAGWCNPPQOHFDLSQPYFORIAQYRAGIVP 122

68 NDGKWCPLPGSIVYATNFPCPNNALPNNAGWCNPPQOHFDLSQPYFORIAQYRAGIVP 127

123 VAYRRVPCVRKRGIRFTINGHSYFNLVLTITNVGAGDVHSAMVKGSRGTGQAMSRNMGON 182

128 ILTYRVPCLRKRGIRFTINGHSYFNLVLTITNVGAGDVHSAMVKGSRGTGQAMSRNMGON 187

183 WQSNSTYLNQGSLSFKYVTSDDGTTVSNXNANAGSFGOTFTGCAHV 227

188 WQSNSTYLNQGSLSFKYVTSDDGTTVSNXNANAGSFGOTFTGCAHV 232

RESULT 4

109818

expansin (clone pteexp2) - loblolly pine (fragment)

C:Species: Pinus taeda (loblolly pine)

C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 26-May-2000

C:Accession: T09818

R:Hutchinson, K.W.; Singer, P.B.; Diaz-Sala, C.; Greenwood, M.S.

A:Description: Expansins are conserved in conifers and expressed in response to exoge

A:Reference number: Z16866

A:Accession: T09818

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-232 <HUT>

A:Cross-references: EMBL:U64890; NID:g1778098; PID:g1778099

A:Experimental source: clone pteexp2; hypocotyl

C:Superfamily: expansin

Query Match 82.8%; Score 1042; DB 2; Length 232;

Best Local Similarity 79.1%; Pred. No. 4.5e-83;

Matches 178; Conservative 25; Mismatches 22; Indels 0; Gaps 0;

3 GGNVNAHTFYGGDASGTMGACGYGNTLSQGYGTNTAALSTALFNNGISCGACFEIRCO 62

8 GGNVNAHTFYGGDASGTMGACGYGNTLSQGYGTNTAALSTALFNNGISCGACFEIRCO 67

63 QNDGKWCPLPGSIVYATNFPCPNNALPNNAGWCNPPQOHFDLSQPYFORIAQYRAGIVP 122

68 NDGKWCPLPGSIVYATNFPCPNNALPNNAGWCNPPQOHFDLSQPYFORIAQYRAGIVP 127

123 VAYRRVPCVRKRGIRFTINGHSYFNLVLTITNVGAGDVHSAMVKGSRGTGQAMSRNMGON 182

128 ILTYRVPCLRKRGIRFTINGHSYFNLVLTITNVGAGDVHSAMVKGSRGTGQAMSRNMGON 187

183 WQSNSTYLNQGSLSFKYVTSDDGTTVSNXNANAGSFGOTFTGCAHV 227

188 WQSNSTYLNQGSLSFKYVTSDDGTTVSNXNANAGSFGOTFTGCAHV 232

RESULT 5

109826

expansin (clone pteexp5) - loblolly pine (fragment)

C:Species: Pinus taeda (loblolly pine)

C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 26-May-2000

C:Accession: T09826

R:Hutchinson, K.W.; Singer, P.B.; Diaz-Sala, C.; Greenwood, M.S.

A:Description: Expansins are conserved in conifers and expressed in response to exoge

A:Reference number: Z16866

A:Accession: T09826

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-232 <HUT>

A:Cross-references: EMBL:U64893; NID:g1778104; PID:g1778105

A:Experimental source: clone pteexp5

C:Superfamily: expansin

Query Match 82.3%; Score 1036; DB 2; Length 232;

Best Local Similarity 78.2%; Pred. No. 1.5e-82;

Matches 176; Conservative 28; Mismatches 21; Indels 0; Gaps 0;

3 GGNVNAHTFYGGDASGTMGACGYGNTLSQGYGTNTAALSTALFNNGISCGACFEIRCO 62

8 GGNVNAHTFYGGDASGTMGACGYGNTLSQGYGTNTAALSTALFNNGISCGACFEIRCO 67

63 QNDGKWCPLPGSIVYATNFPCPNNALPNNAGWCNPPQOHFDLSQPYFORIAQYRAGIVP 122

68 NDGKWCPLPGSIVYATNFPCPNNALPNNAGWCNPPQOHFDLSQPYFORIAQYRAGIVP 127

123 VAYRRVPCVRKRGIRFTINGHSYFNLVLTITNVGAGDVHSAMVKGSRGTGQAMSRNMGON 182

128 ILTYRVPCLRKRGIRFTINGHSYFNLVLTITNVGAGDVHSAMVKGSRGTGQAMSRNMGON 187

Genome version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 11, 2002, 14:53:05 ; Search time 11.4421 Seconds
(without alignments)
1914.721 Million cell updates/sec

Title: US-09-896-301-2

Perfect score: 1259
Sequence: 1 AGGGMVNAHATFYGGGDSG.....NNXANAGWSFGQFTGAHVR 228

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1246	99.0	237	2	T50654 expansin EXP1 [imp
2	1115	88.6	248	2	C84444 probable expansin
3	1047	83.2	232	2	T09821 expansin (clone pp
4	1042	82.8	232	2	T09818 expansin (clone pp
5	1036	82.3	232	2	T09826 expansin (clone pp
6	1028	81.7	232	2	T09825 expansin (clone pp
7	1026	81.5	250	2	T10079 expansin S1 precu
8	1001	79.5	246	2	T04175 expansin 2 - rice
9	979	77.8	251	2	T03298 expansin 2 - rice
10	976.5	77.6	253	2	F84831 probable expansin
11	968	76.9	258	2	T09786 expansin - upland
12	955	75.9	257	2	D84820 probable expansin
13	936	74.3	257	2	T50658 expansin 9 [import
14	930.5	73.9	255	2	T50655 expansin EXP5 [imp
15	929	73.8	260	2	T47689 expansin-like prot
16	921	73.0	262	2	T02530 probable expansin
17	919.5	73.0	255	2	T50656 expansin EXP2 [imp
18	917	72.8	262	2	T50660 alpha-expansin 2 [
19	916	72.8	258	2	S53082 pollen allergen ho
20	913	72.5	258	2	T48247 expansin 1 - toma
21	910	72.3	261	2	T07630 alpha-expansin Ose
22	910	72.3	264	2	T50659 probable expansin
23	908	72.1	257	2	T02127 expansin EXP6 [imp
24	908	72.1	259	2	T50653 expansin 18 - toma
25	898	71.3	255	2	T06573 probable expansin
26	889	70.6	260	2	T100816 expansin S2 precu
27	879	69.8	258	2	T10083 expansin - rice
28	863.5	68.6	261	2	T03737 hypothetical prote
29	832.5	66.1	252	2	F86335

30	784	62.3	255	2	T02010 expansin homolog T
31	780.5	62.0	255	2	T03299 expansin 3 - rice
32	670	53.2	257	2	G96654 hypothetical prote
33	662	52.6	160	2	T09871 expansin - upland
34	662	52.6	257	2	F86259 protein T1C24.10
35	599	47.6	256	2	T05648 expansin homolog F
36	318.5	25.3	102	2	T09828 expansin (clone pp
37	302	24.0	77	2	T09815 expansin (clone pd
38	287	22.8	277	2	S48032 cimi protein - soy
39	282	22.4	81	2	T09041 cimi protein homol
40	274	21.8	86	2	T09830 expansin (clone pp
41	274	21.8	259	2	T50657 beta-expansin [imp
42	274	21.8	271	2	H84592 beta-expansin [imp
43	271	21.5	261	2	T04301 beta-expansin - ri
44	256.5	20.4	491	2	F96681 protein F1E22.6 [i
45	230.5	18.3	259	2	F84886 probable beta-expa

ALIGNMENTS

```

RESULT 1
T50654
expansin EXP1 [imported] - Arabidopsis thaliana (fragment)
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 02-Sep-2000
C:Accession: T50654
R:Shcherban, T.Y.; Shi, J.; Durachko, D.M.; Gultinan, M.J.; McQueen-Mason, S.J.; Shi
Proc. Natl. Acad. Sci. U.S.A. 92, 9245-9249, 1995
A:Title: Molecular cloning and sequence analysis of expansins--a highly conserved, mu
A:Reference number: Z14894; MUID:96016146
A:Accession: T50654
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-237 <SHC>
A:Cross-references: EMBL:U30476; PIDN:AA838070.1
A:Genetics:
A:Gene: EXP1
C:Function:
A:Description: induces extension (creep) in plant cell walls
C:Superfamily: expansin
C:Keywords: cell wall

Query Match          99.0%  Score 1246; DB 2; Length 237;
Best Local Similarity 98.7%; Pred. No. 9.4e-101;
Matches 225; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGGMVNAHATFYGGGDSGTMGACGYGNLYSGYGTNTALSTALFNNGISGACFEI 60
    |||
DB 10 AGGGMVNAHATFYGGGDSGTMGACGYGNLYSGYGTNTALSTALFNNGISGACFEI 69
    |||

QY 61 KCQNDGKWCPLPGSIYVATNFCPPNNALPNNAGCMCNPCPOHFDLSQYFQRIADYRAGI 120
    |||
DB 70 KCQNDGKWCPLPGSIYVATNFCPPNNALPNNAGCMCNPCPOHFDLSQYFQRIADYRAGI 129
    |||

QY 121 VVAVARVPCVRGGRFTINGHSYFNLITNVGAGDVHSAMKGSRTGQASRMWG 180
    |||
DB 130 VVAVARVPCVRGGRFTINGHSYFNLITNVGAGDVHSAMKGSRTGQASRMWG 189
    |||

QY 181 QNMQSNLYNGQSLSEFKVTTSDQGTIVSNXNAGMSEFGQFTGAHVR 228
    |||
DB 190 QNMQSNLYNGQSLSEFKVTTSDQGTIVSNXNAGMSEFGQFTGAHVR 237
    |||

RESULT 2
C84444
probable expansin [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C:Accession: C84444
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanKen, S.E.; Umayam, L.; Tallon,
euss, D.; Niernan, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter

```


DB 221 GSGGALYSDDGIDDIDONAVYLFRFENALTTAIGKGAVC-----CLPTSGS 266
 QY 76 VTATNCPNNALPNNAGCNCPPQHPDLSPVFORIAGYRAGIYPVAVRVPVCRRCG 135
 DB 267 STTV-----PIVTFSDKKQLVFERNNHSMKG--GAIYARRLSTSSGG 306
 QY 136 IRETINGHSYFNLVLTNNGA-----GPHSAMYKSGRTQWAMSRMGQWMOGNSYL 189
 DB 307 PLTFINNISYAN---SQNLGGALAIPTGGEISLSAKKGTIT-----FGGNRTSLPPL 355
 QY 190 NG---QSLSF-KVTTSDQTI 206
 DB 356 NGIHLDMNAKFKLKQANGYSI 377

RESULT 14
 YMS5_CAEEL STANDARD; PRT; 1385 AA.

ID YMS5_CAEEL STANDARD; PRT; 1385 AA.
 AC P34501;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Hypothetical 159.2 KDa protein K03H1.5 in chromosome III.
 GN K03H1.5.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
 RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
 RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
 RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
 RA Sims M., Smalton N., Smith A., Smith M., Sonhammer E., Staden K.,
 RA Sulston J., Thierry-Mieg J., Thomas K., Vaubin M., Vaughan K.,
 RA Watson R., Watson A., Weinstock L., Wilkinson-Sprat J.,
 RA Wohldman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RA elegans.";
 RL Nature 368:32-38(1994).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; Z29560; CAA82664.1; -;
 DR PIR; S41028; S41028.
 DR WormPep; K03H1.5; CE03459.
 DR InterPro; IPR002909; IPT_TIG.
 DR InterPro; IPR003886; Nidogen_ext.
 DR InterPro; IPR000436; Sush1_SCR_CCP.
 DR InterPro; IPR001846; Vwd.
 DR Pfam; PF00084; sush1; 1.
 DR Pfam; PF00094; vwd; 2.
 DR SMART; SM00032; CCP; 1.
 DR SMART; SM00429; IPT; 1.
 DR SMART; SM00539; NIDO; 1.
 DR SMART; SM00216; VWD; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 1385 AA; 159181 MW; BDCD8F59CEA38C03 CRC64;

Query Match 6.9%; Score 86.5; DB 1; Length 1385;
 Best Local Similarity 22.4%; Pred. No. 9.2;

Matches 50; Conservative 24; Mismatches 76; Indels 73; Gaps 10;
 QY 1 AGG-----GNNVNAHAFYGGGDSGTMG-GACGYGLYSOGYNTNTALSTALFNNGISCG 55
 DB 393 AGGLDGFGRQAKAMAFNGNCTGMYGLPYSGEGRMLKGY-----FSNVLP-- 439
 QY 56 ACEFTRQNDGW-----CLPGSI--VTATNCPNNALPNNAGCNCPE-- 99
 DB 440 -----PGWTHRVDEVITIPACCTNANSGNMFTAPWPGFMHGMALINSGPCLRAD 490
 QY 100 --QGFHDSQVFORIADYRAGIYPVAVRVPV-----RRCGIRETINGHSYFNLVLT 152
 DB 491 SVKVNENMOTGSTRLSRVBARCIMPFMFKIGLIVIRMSRQGSFPPEKFFV----- 544
 QY 153 NVGAGDVHSAVYKGS-----RTGQAMSRW 179
 DB 545 -VNSERAPASVSLKSDVDNKNRWYEPYAQELALGQAMNLTW 586

RESULT 15
 YF48_MYCTU STANDARD; PRT; 678 AA.

ID YF48_MYCTU STANDARD; PRT; 678 AA.
 AC Q10778;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical PPE-family protein RV1548C.
 GN RV1548C OR MT1599 OR MYCT48.17.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-H37RV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence.";
 RL Nature 393:537-544(1998).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; Z74020; CAA96335.1; -;
 DR EMBL; AE007026; AAK45866.1; ALT_INIT.
 DR TIGR; MT1599; -;
 DR Tuberculist; RV1548C; -;
 DR InterPro; IPR002989; Mycobact_pentapep.

FT CHAIN 113 333 CATHEPSIN J.
 FT ACT_SITE 137 137 BY SIMILARITY.
 FT ACT_SITE 275 275 BY SIMILARITY.
 FT ACT_SITE 299 299 BY SIMILARITY.
 FT DISULFID 134 177 BY SIMILARITY.
 FT DISULFID 168 210 BY SIMILARITY.
 FT DISULFID 268 321 BY SIMILARITY.
 FT CARBOHYD 71 71 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 216 216 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 220 220 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 267 267 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 42 42 P -> PK (IN REF. 2).
 SQ SEQUENCE 333 AA; 37147 MW; 79A8FFID5A13B721 CRC64;

Query Match 7.4%; Score 93; DB 1; Length 333;
 Best Local Similarity 24.2%; Pred. No. 0.6;

Matches 45; Conservative 15; Mismatches 72; Indels 54; Gaps 8;

QY 17 DASGTMGAGCGYGNLSQGYGTNTA--ALSTALFNNGLSGACF-----EIRCOND 65
 DB 167 DCSKTYG-----NKGQSGTAHQAFYVYLNKNGLEATPYPEKGDGPRYRSEN- 216
 QY 66 GKACGLGSLVYVATNFCPPNNALPNNAGCNCPPQGHFDLSQPFQRIAYRAGIVPAY 125
 DB 217 -----ASANITDYVNLPPNELYLVAVASIGFVSAIDASHDSFR----- 256
 QY 126 RRVPCVRGRIPTINGHSYF--NLVLTINVGAGDVHSAMVYSGRTGQAMRNMGOMW 183
 DB 257 -----FYNGCITYEPNCSTFVNHAVLVVYGSEGD-----YADGNVYV-LINKSWEEM 305
 QY 184 QSNSTL 189
 DB 306 GNMGYM 311

RESULT 11
 ROH2_HUMAN STANDARD; PRT; 449 AA.

AC P55795; Q9HHA7;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Heterogeneous nuclear ribonucleoprotein H' (hnRNP H') (FMP-3).
 GN HNRPH2 OR FMP3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95048329; PubMed=7959728;
 RA Vorechovsky I., Vetric D., Holland J., Bentley D.R., Thomas K.,
 RA Zhou J.N., Notarangelo L.D., Plebani A., Fontan G., Ochs H.D.;
 RT "Isolation of cosmid and cDNA clones in the region surrounding the
 RT BTK gene at Xq21.3-q22.";
 RL Genomics 21:517-524(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Oeljen J.C., Liu X., Lu J., Malley T.M., Allen R.C., Muzny D.M.,
 RA Belmont J.W., Gibbs R.A.;
 RT Submitted (NOV-1996) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP PARTIAL SEQUENCE FROM N.A.
 RX MEDLINE=96081943; PubMed=7499401;
 RA Honore B., Rasmussen H.H., Vorum H., Dejgaard K., Liu X.,
 RA Gronov P., Madsen P., Gesser B., Tommerup N., Celis J.E.;
 RT "Heterogeneous nuclear ribonucleoproteins H', H', and F are members of
 RT a ubiquitously expressed subfamily of related but distinct proteins
 RT encoded by genes mapping to different chromosomes.";
 RL J. Biol. Chem. 270:28780-28789(1995).

CC -i- FUNCTIONAL, TRANSLATABLE MRNAS IN THE CYTOPLASM. BINDS POLY (RG).
 CC -i- SUBCELLULAR LOCATION: Nuclear; nucleoplasm.
 CC -i- TISSUE SPECIFICITY: EXPRESSED UBQUITOUSLY.
 CC -i- SIMILARITY: CONTAINS 3 RNA RECOGNITION MOTIFS (RRM).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U01923; NOT ANNOTATED_CDS.
 DR EMBL: U78027; AAB64202.1; -.
 DR Aairus/Ghent-2DPAGE: 4432; IEF.
 DR MIM: 601036; -.
 DR InterPro: IPR000504; RRM.
 DR Pfam: PF00076; RRM: 3.
 DR SMART: SM00360; RRM: 3.
 DR PROSITE: PS0102; RRM: 3.
 DR PROSITE: PS0030; RRM_RNP_1; FALSE_NEG.
 KW Nuclear protein; RNA-binding; Ribonucleoprotein; Repeat.
 FT DOMAIN 11 90 RNA-BINDING (RRM) 1.
 FT DOMAIN 111 188 RNA-BINDING (RRM) 2.
 FT DOMAIN 289 364 RNA-BINDING (RRM) 3.
 FT REPEAT 234 433 2 X 16 AA GLY-RICH APPROXIMATE REPEATS.
 FT REPEAT 234 249 1-1.
 FT REPEAT 418 433 1-2.
 FT DOMAIN 354 392 2 X 19 AA PERFECT REPEATS.
 FT REPEAT 354 372 2-1.
 FT REPEAT 374 392 2-2.
 SQ SEQUENCE 449 AA; 49263 MW; C892523A638F07C7 CRC64;

Query Match 7.1%; Score 89; DB 1; Length 449;
 Best Local Similarity 20.3%; Pred. No. 1.8;

Matches 48; Conservative 28; Mismatches 113; Indels 48; Gaps 8;

QY 2 GGGWNAHATPYGGDASGTMGAGCGYGNLSQGYGTNTALSTALFNNGLSGACFETR 61
 DB 225 GAGFERMRGAGGAGG-----YGGYDYG--YNDGFGFSDFR-----GIDLNYCF5-- 269
 QY 62 CONDKWCLPGSLVYVATNFCPPNNALPNNAG-----GMCNPPQGHFDL-----SQ 107
 DB 270 GMSDHRVGDGSSPSTGHCYHMGRLPYRATENDIYFSPPLNMRHIEIGDGRVTC 329
 QY 108 PVFQRIAYRAGIVPAYRVPCVRR-----GGIRPTINGHSYFNVLITNVGAG 158
 DB 330 EADVEFATHEDAVAMAKDKAMQHRVYELFLNSTAGTSGGAYDSHYELFLNSTAGASG 389
 QY 159 DVHSAWVNGSR-----TGQAMSRNMGQNMNSYINGOSLSKVTTSQGT 205
 DB 390 GAYGOMMGGMGLSNSSYGGPASPQSLSGYGGYGGSSMGYDQVLQENSSDYQS 446

RESULT 12

YCC8_YEAST STANDARD; PRT; 405 AA.

AC P25367;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Hypothetical 42.6 kDa protein in Birk1-FUS1 intergenic region.
 GN YCLO28W OR YCL28W OR YCL181.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;

RP SEQUENCE FROM N.A.
 RX MEDLINE=91377317; PubMed=1897318;
 RA Rad M.R., Luetzenkirchen K., Xu G., Kleinhans U., Hollenberg C.P.;
 RT "The complete sequence of a 11,953 bp fragment from C1G on chromosome

Db 67 ASWCGAGCGKCYLTSTGQAPCCSGTGAAGSGIIVMTNLCPPN-----GNA-QWC--P 119
 Qy 100 -----QOHPDL 105
 |||:
 Db 120 VVGSTNOGYGYVHPDI 135

RESULT 9
 ID PER_DRONE STANDARD; PRT; 385 AA.
 AC P91686;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Period circadian protein (Fragment).
 GN PER.
 OS Drosophila nebulosa (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7271;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=0761.0 PALMIRA / COLOMBIA;
 RX MEDLINE=97357421; PubMed=9214747;
 RA Gleason J.M., Powell J.R.;
 RT "Interspecific and intraspecific comparisons of the period locus in
 the Drosophila willistonii sibling species.";
 RL Mol. Biol. Evol. 14:741-753(1997).
 CC -1- FUNCTION: ESSENTIAL FOR BIOLOGICAL CLOCK FUNCTIONS. DETERMINES THE
 PERIOD LENGTH OF CIRCADIAN AND ULTRADIAN RHYTHMS; AN INCREASE IN
 PER DOSAGE LEADS TO SHORTENED CIRCADIAN RHYTHMS AND A DECREASE
 LEADS TO LENGTHENED CIRCADIAN RHYTHMS. ESSENTIAL FOR THE CIRCADIAN
 RHYTHMICITY OF LOCOMOTOR ACTIVITY, ECLOSION BEHAVIOR, AND FOR THE
 RHYTHMIC COMPONENT OF THE MALE COURTSHIP SONG THAT ORIGINATES IN
 THE THORACIC NERVOUS SYSTEM. THE BIOLOGICAL CYCLE DEPENDS ON THE
 RHYTHMIC FORMATION AND NUCLEAR LOCALIZATION OF THE TIM-PER
 COMPLEX. LIGHT INDICES THE DEGRADATION OF TIM, WHICH PROMOTES
 ELIMINATION OF PER. NUCLEAR ACTIVITY OF THE HETERODIMER
 COORDINATIVELY REGULATES PER AND TIM TRANSCRIPTION THROUGH A
 NEGATIVE FEEDBACK LOOP. BEHAVES AS A NEGATIVE ELEMENT IN CIRCADIAN
 TRANSCRIPTIONAL LOOP. DOES NOT APPEAR TO BIND DNA, SUGGESTING
 INDIRECT TRANSCRIPTIONAL INHIBITION (BY SIMILARITY).
 CC -1- SUBUNIT: FORMS HETERODIMER WITH TIMELESS (TIM); THE COMPLEX THEN
 SUBLOCALIZES INTO THE NUCLEUS (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR AT SPECIFIC PERIODS OF THE DAY.
 CC FIRST ACCUMULATES IN THE PERINUCLEAR REGION ABOUT ONE HOUR BEFORE
 TRANSLOCATION INTO THE NUCLEUS. INTERACTION WITH TIM IS REQUIRED
 FOR NUCLEAR LOCALIZATION (BY SIMILARITY).
 CC -1- PTM: PHOSPHORYLATED WITH A CIRCADIAN RHYTHMICITY, PROBABLY BY THE
 DOUBLE-TIME PROTEIN (DBT). PHOSPHORYLATION COULD BE IMPLICATED IN
 THE STABILITY OF PER MONOMER AND IN THE FORMATION OF HETERODIMER
 PER-TIM (BY SIMILARITY).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U51090; AAB41386.1; -;
 DR FLYBase; FBgn0018750; Dnelyper.
 KW Biological rhythms; Repeat; Nuclear protein; Phosphorylation.
 FT NON_TER 1
 FT DOMAIN 102 112 POLY-GLY.
 FT DOMAIN 213 218 POLY-SER.
 FT DOMAIN 226 234 POLY-GLY.
 FT DOMAIN 246 252 POLY-GLY.
 FT DOMAIN 377 383 POLY-ALA.
 FT NON_TER 385 385
 SQ SEQUENCE 385 AA; 39104 MW; A6C519E2CE66D5F0 CRC64;

Query Match 7.5%; Score 94.5; DB 1; Length 385;
 Best Local Similarity 26.3%; Pred. No. 0.52;
 Matches 41; Conservative 18; Mismatches 46; Indels 51; Gaps 9;

Qy 1 AGGQWYNAHATFYGGDASGTMGACGCGYGNLYSQYGTATLALSTALF-NGLSCGACPE 59
 |||:
 Db 225 AGG-----GGGGGASGSGTGTGNNAGGGGGSMAOSSTMYQSGLSG----- 269
 Qy 60 IRCQNDGKN-----CLPGSIV-----VTATNFCPPNALDN-----NAGC 94
 |||:
 Db 270 --TQNNINLPSPSVGITTPTTSVLTSMAYVAQSEFSQHSFPFTFYITPASIASSPAST 327
 |||:
 Qy 95 WCPN-POQHPDL-----SOPVFORIAQYRAGIYPVAY 125
 |||:
 Db 328 SPNRPKHHTVHPSEOP---STSGGAATMPLQY 360

RESULT 10
 ID CATL_MOUSE STANDARD; PRT; 333 AA.
 AC Q9R014; Q9WV51;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cathepsin J precursor (EC 3.4.22.-) (Cathepsin P) (CatLrp-P).
 GN CTSJ OR CTSJ.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Embryo;
 RX MEDLINE=9456833; PubMed=10526153;
 RA Tislar K., Deussing J., Peters C.;
 RT "Cathepsin J, a novel murine cysteine protease of the papain family
 with a placenta-restricted expression.";
 RL FEBS Lett. 459:299-304(1999).
 CC [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Placenta;
 RA Sol-Church K., French J., Troeber D., Mason R.W.;
 RT "Cloning of a mouse cysteine protease.";
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Lysosomal (Potential).
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN PLACENTA.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1; ALSO KNOWN AS THE
 PAPAIN FAMILY OF THIOL PROTEASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF136272; AAF13142.1; -;
 DR EMBL; AF158182; AAD41898.1; -;
 DR HSSP; P07711; 1CUL.
 DR MEROPS; C01.038; -;
 DR MGD; MGI:1349426; Ctsj.
 DR InterPro; IPR000668; Peptidase_C1.
 DR InterPro; IPR000169; ThiolProt_act_site.
 DR Pfam; PF00112; Peptidase_C1; 1.
 DR PRINTS; PR00705; PAPAIN.
 DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
 DR PROSITE; PS00639; THIOL_PROTEASE_HIS; 1.
 DR PROSITE; PS00640; THIOL_PROTEASE_ASN; 1.
 DR Hydrolase; Thiol protease; Glycoprotein; Lysosome; Zymogen; Signal.
 FT SIGNAL 1 17
 FT PROPEP 18 112 ACTIVATION PEPTIDE.


```

OS Phalaris aquatica (Canary grass)
OC Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Poaceae; Phalaris.
OX NCBI_TaxID=28479;
RN RP
RC SEQUENCE FROM N.A.
RC TISSUE=Pollen;
RX MEDLINE=96105569; PubMed=8564724;
RA Suphioglu C., Singh M.B.;
RT "Cloning, sequencing and expression in Escherichia coli of pha 1
RT and four isoforms of pha 5, the major allergens of canary grass
RT pollen.";
RL Clin. Exp. Allergy 25:853-865(1995).
RN [2]
RP SEQUENCE OF 30-49.
RX MEDLINE=93131901; PubMed=7687099;
RA Suphioglu C., Singh M.B., Simpson R.J., Ward L., Knox R.B.;
RT "Identification of canary grass (Phalaris aquatica) pollen allergens
RT by immunoblotting: IgE and IgG antibody-binding studies.";
RL Allergy 48:273-281(1993).
CC -1- SUBCELLULAR LOCATION: secreted.
CC -1- DISEASE: CAUSES GRASS POLLEN ALLERGY.
CC -1- SIMILARITY: BELONGS TO THE LIL P I FAMILY OF ALLEGENS.
CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE ECG5 DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE CBD DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: S80654; AAB35984.1; -.
CC HSP: P43214; IWHO.
DR InterPro: IPR000882; Pollen_allergen.
DR Pfam: PF01357; Pollen_allergen: 1.
DR PRINTS: PRO1225; EXPANSINFAMLY.
DR ProDom: PD002179; Pollen_allergy: 1.
DR PROSITE: PS50843; EXPANSIN_CBD: 1.
DR PROSITE: PS50842; EXPANSIN_ECG5; 1.
KW Allergen; Glycoprotein; Signal.
FT SIGNAL 1 29
FT CHAIN 30 269 MAJOR POLLEN ALLERGEN PHA A 1.
FT DOMAIN 67 173 EXPANSIN-LIKE ECG45.
FT DOMAIN 187 268 EXPANSIN-LIKE CBD.
FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT COMFLICT 37 37 P -> G (IN REF. 2).
SQ SEQUENCE 269 AA; 29011 MM; 05883458ACEB77F CRC64;
Query Match 15.8%; Score 199; DB 1; Length 269;
Best Local Similarity 26.8%; Pred. No. 4.8e-10;
Matches 62; Conservatly 39; Mismatches 100; Indels 30; Gaps 9.
OY 2 GCGWYNNAHATFPGGCGDAGTGM--GGACGYGNLYSOGYGTNTAALSTALFNNGLSGCAGE 59
Dy 44 GDKWLDASTWYGRKGTAGPKNDGACGCKVDKAPFNMTGCGMTPIPKDRGCGSGCE 103
OY 60 IRCQNDGKWCILPGSTIVTATNFCPPNNALPNNAGGCMCPDQHFDSLQPFORLQD---- 115
Dy 104 LKCSKP-ESCGSEPTVHTD-----DNEEP-----IAPHFDLSGHAFGSMKKEE 150
OY 116 ---YAGIVPAVYRVRPCVRRGGINFTIN---GHSYFNLYLTWVGAGDVHSAVVK-G 167
Dy 151 ENVRGAGELEQERFVKKYKCPDGTPTTHVEKGSNPNYALLVKKVGGDVAVADIKER 210
OY 168 SRTGQASNRNMGWMSNS--LYLNGQSLSFKVTTSDCQTIYSNNXANAGW 216
Dy 211 GKDKKIEIKESMGALWIRIDTPDKLKG-PTTVATTEGCGTAKAEFDYIPGHW 260

```

ID	MPOL_ORYSA	STANDARD:	PRT:	263 AA.
AC	040638:			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	01-MAR-2002 (Rel. 41, Last annotation update)			
DE	Major pollen allergen Ory s 1 precursor (Ory s I).			
OS	Oryza sativa (Rice).			
OC	Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;			
OC	Ehmerioideae; Oryzaceae; Oryza.			
OX	NCHI_TaxID=4530;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Anther;			
RX	MEDLINE=96069591; PubMed=7590339;			
RA	Xu H., Theerakulpisut P., Gouding N., Suphloglu C., Singh M.B.,			
RA	Bhalla P.L.;			
RT	"Cloning, expression and immunological characterization of Ory s 1,			
RT	the major allergen of rice pollen.";			
RL	Gene 164:255-259(1995).			
CC	- 1- SUBCELLULAR LOCATION: Secreted.			
CC	- 1- TISSUE SPECIFICITY: EXPRESSED IN MATURE ANTHEMS BUT NOT IN			
CC	VEGETATIVE OR OTHER FLORAL TISSUES.			
CC	- 1- DISEASE: CAUSES GRASS POLLEN ALLERGY.			
CC	- 1- SIMILARITY: BELONGS TO THE LOT P I FAMILY OF ALLERGENS.			
CC	- 1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE EG45 DOMAIN.			
CC	- 1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE CBD DOMAIN.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see http://www.isb-slb.ch/announce/			
CC	or send an email to license@isb-slb.ch).			
CC	-----			
DR	EMBL: U31771; AAA86533.1; .			
DR	HSSP: P43214; IMHO.			
DR	InterPro: IPR000882; Pollen_allergen.			
DR	Pfam: PF01357; Pollen_allergen. 1.			
DR	PRINTS: PRO1225; EXPANSINFMALY.			
DR	ProDom: PD002179; Pollen_allergen. 1.			
DR	PROSITE: PS50843; EXPANSIN_CBD; 1.			
DR	PROSITE: PS50842; EXPANSIN_EG45; 1.			
KW	Allergen; Glycoprotein; Signal; Multigene family.			
FT	SIGNAL 1 23			POTENTIAL.
FT	CHAIN 24 263			MAJOR POLLEN ALLERGEN Ory s 1.
FT	DOMAIN 61 164			EXPANSIN-LIKE EG45.
FT	DOMAIN 178 259			EXPANSIN-LIKE CBD.
FT	CAROHSD 32 32			N-LINKED (GLCNAC. . .) (POTENTIAL).
SO	SEQUENCE 263 AA; 28497 MW; BIC5F24EA39BD60 CRC64;			
QY	Query Match 13.4%;			Score 168.5; DB 1; Length 263;
	Best Local Similarity 29.5%;			Pred. No. 1.8e-07;
	Matches 57; Conservative 29; Mismatches 74; Indels 33; Gaps 11.			
QY	23 GGAGCGNLYSGY-GTNALSTLFLNNGLSCGCFEIRCONDGKWC-LPGSYIVYNT 79			
DB	61 GGAGCGKDDVDAKPELGMNSCG-NDRPEKGGKCGSCFETGCKR-EGCDKPLALHVTMD 118			
QY	80 NCEPNNALPNNAGCMCNPPQHPDLSCQVPEFORIQ---YAGIVPVVYRRPVCVRIGC 135			
DB	119 NDEP-----IAAYHFDLSGLAMAKDKGDELKRAAGIIDIQFFRRVKKCKPYAD 164			
QY	136 IFFTN---GHSYENLVLTIVNGAGADVHSAV--KSGRTGQWASRWGQNNWOSN--- 187			
DB	165 TTFTHIEKASNPMLTALLVKKVYAGDGDVVEVEIKEGGSEE-WKALKESGMAIMRIDTPK 223			
QY	188 YLNGOSLSFKVTT 200			
DB	224 PLKG-PFSVRVTT 235			

```

RT group Allergens from eight different grass species.";
RL J. Allergy Clin. Immunol. 94:689-698(1994).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE LOL P 1 FAMILY OF ALLERGENS.
CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE EG45 DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE CBD DOMAIN.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).
-----
CC EMBL; X78813; CA55390.1; -.
CC HSSP; P43214; IMHO.
CC InterPro; IPR000882; Pollen_allergen.
CC Pfam; PF01357; Pollen_allergen; 1.
CC PRINTS; PRO1225; EXPANSINFAMILY.
CC ProDom; PD002179; Pollen_allergen; 1.
CC PROSITE; PS50843; EXPANSIN_CBD; 1.
CC PROSITE; PS50842; EXPANSIN_EG45; 1.
CC Allergen; Glycoprotein; Signal; Multigene family.
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 263 POLLEN ALLERGEN PHL P 1.
FT DOMAIN 61 167 EXPANSIN-LIKE EG45.
FT DOMAIN 181 262 EXPANSIN-LIKE CBD.
FT CARBOHYD 32 32 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 263 AA; 28457 MW; 046BA249C17BC048 CRC64;

Query Match 16.7%; Score 210; DB 1; Length 263;
Matches Local Similarity 27.3%; Pred. No. 5.5e-11;
Matches 63; Conservative 40; Mismatches 98; Indels 30; Gaps 9;

QY 2 GGGWNAHATPFGGADASTM--GGAGCGGNLYSGYGTNTALSTALENNLSGACPE 59
Db 38 GDKWLDAKSTWYGRKTGAGPKDNGACGCKYDKRPFSGMTGCGNTPIFKSRGGSCFE 97
QY 60 IRCQNDKWCMLPGSLIVVATNFCPPNNALPNNAGCWCNPPQOHFDLSQVFQRIQA--- 115
Db 98 IKCTKP-EACGSEPPVVAHTD---DNEP-----IAPHFDLSGIAFPAMAKKQDE 144
QY 116 ---TRAGIVPAVARYRVCVRRCGRIGFTTN---GISTENLVLTNVGAGADVHSMVK G 167
Db 145 QKLRSAGLELEDFRRKCKYPEGTGVTFTFVHEKGSNPNTALALVKKYVNGDGDAVADIEK 204
QY 168 SRTGQASRMKQWQMSNS--YLNGQSLSEFVTN SDQTIYSNNXANAGW 216
Db 205 GKDKWIELKESGAIWRIDTPDKLTG-PFTVATYTGEGTKTEADVDYIPEGW 254

RESULT 4
MPHL_HOLLA STANDARD: PRT: 265 AA.
AC P43216; Q39975;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Major pollen allergen Hol 1 I precursor (Hol 1 I) (Hol 1 1.0101 and
DE 1.0102).
OS Holcus lanatus (Velvet grass).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Poae; Holcus.
OX NCBI_Taxid=29679;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. AVENEAE; TISSUE=Pollen;
RA Schramm G.-D., Bufer A., Becker W.M., Schlaak M.;
RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
RP [2]
SQ SEQUENCE OF 18-265 FROM N.A.

```

```

RC STRAIN=CV_AVENENAE; TISSUE=Pollen;
RA MEDLINE=97358126; PubMed=9215246;
RX Schramm G.D., Bute A., Petersen A., Haas H., Schlaak M., Becker W.M.;
RT "Mapping of IgE-binding epitopes on the recombinant major group I
RT allergen of velvet grass pollen, rHol 1.1.";
RL J. Allergy Clin. Immunol. 99:781-787(1997).
RN [3]
RP CHARACTERIZATION.
RC STRAIN=CV_AVENENAE; TISSUE=Pollen;
RX MEDLINE=96319506; PubMed=8768803;
RA Schramm G.D., Petersen A., Bute A., Schlaak M., Becker W.M.;
RT "Identification and characterization of the major allergens of velvet
RT grass (Holcus lanatus), Hol 1.1 and Hol 1.5.";
RN Int. Arch. Allergy Immunol. 110:354-363(1996).
CC -I- SUBCELLULAR LOCATION: Secreted.
CC -I- SIMILARITY: BELONGS TO THE LOL P I FAMILY OF ALLEGENS.
CC -I- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE EG45 DOMAIN.
CC -I- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE CBD DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; 227084; CAA81610.1; -.
DR EMBL; 268893; CAA93121.1; -.
DR HSSP; P43214; IMHO.
DR InterPro: IPR000882; Pollen_allergen.
DR Pfam; PF03357; Pollen_allergen; 1.
DR PRINTS; PRO1225; EXPANSINFAMLY.
DR PRODOM; PD002179; Pollen_allergen; 1.
DR PROSITE; PS50843; EXPANSIN_CBD; 1.
DR PROSITE; PS50842; EXPANSIN_EG45; 1.
KW Allergen; Signal.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 265 MAJOR POLLEN ALLERGEN HOL 1.1.
FT DOMAIN 63 169 EXPANSIN-LIKE EG45.
FT DOMAIN 183 264 EXPANSIN-LIKE CBD.
FT VARIANT 103 103 T -> S (IN HOL 1.1.0102).
SQ SEQUENCE 265 AA; 28590 MW; 1EE23B3EE19BAD6D CXC64;

Query Match 15.8%; Score 199; DB 1; Length 265;
Best Local Similarity 26.4%; Pred. No. 4.7e-10;
Matches 61; Conservative 41; Mismatches 99; Indels 30; Gaps 9;

OY 2 GGGWNAIATPTFGGSDASGTM--GGACGYGNLISQGYGTNTAALSTALFNNNGISGCAFE 59
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 40 GDEWIDASTWYGRKPTGAGPKDNGAGCKVDKDPFGSGMGCGTPTFFKGRGSGSCE 99
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 60 IRCQNDGKMCPLGSLIVYATNCEPCNNMLPNNAGGWCNPDPOHFDLSQPFVORLQAO---- 115
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 100 IKCTPT-ESCSEPTVHTID- -DNEEP-----IAPYHPLDSGHANGSMKKKEE 146
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 116 --YRAGIVPAVYRRPVCVRRGIRFTIN---GHSYENLVLTNVGAGDVHSAVMK-G 167
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 147 QKLRSGAELELKRFRVCKKYPDGTPTFTHVEKGSNPYLLALLVKKYIDGDGDVAVDIKEK 206
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 168 SRTGQAMSRMGWQMSNS--YLNQGSLSFVTTSDQOTIVSNXNAAGW 216
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 207 GKDKWTELKESGWAVRVDTPDKILG-FTVRYTYTEGGTGKGEADVPEGW 256
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 5
ID MPAL_PHAQO STANDARD; PRT; 269 AA.
AC Q41260;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Major pollen allergen Pha a 1 precursor (Pha a I).

```

Db 15 GSKWLEAATFYGSGNPGCAAPDDHGGAGYKVDKPPEDMTAGCEPIPKDGLGCRACY 74
 Qy 59 EIRCONDKKWCPLPGSIVVTATNFCPPNNALPNNAGWCNPPQGHFDLSQPVFORIO--- 115
 Db 75 EIKCKEVE-CSGEPVLYKTID-----KNEHIAATHFDLSGAFAMAKKGG 121
 Qy 116 ---YRAGIVPAVYRRVPCVRRGGIRFTIN---GHSYFNLVLTITVNGAGDVHSAWVG 167
 Db 122 EDKLRKAGELTLOFRVYKCKYPSGKTITFEIKGSNDHYLALVYKVAAGDGNIVAVDIKP 181
 Qy 168 SRIC-WQAMRNNGQNNQ--SNSYLNQSLSFVYTSDDGTIYSNNXANAGSFGQTFT 223
 Db 182 RDSDEFIPKMSWGAIVRIDPKKPLKG-PPSIRLTSEGAHLVQDDVIVANMKPDTVYT 239
 RESULT 2
 MPIL_LOLPR STANDARD; PRT; 263 AA.
 AC P14946; P19964;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Pollen allergen Lol p 1 precursor (Lol p I) (Allergen R7).
 OS Lolium perenne (Perennial ryegrass).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
 OC Poaceae; Lolium.
 OC NCBI_TaxId=4522;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90375479; PubMed=1697854;
 RA Perez M., Ishioka G.Y., Walker L.E., Chesnut R.W.;
 RT "cDNA cloning and immunological characterization of the rye grass
 J. Biol. Chem. 265:16210-16215(1990).
 RL [2]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 24-48.
 RC TISSUE=Pollen;
 RX MEDLINE=91160716; PubMed=2001733;
 RA Griffith I.J., Smith P.M., Pollock J., Theerakulpisut P.,
 RA Avjoglou A., Davies S., Hough T., Singh M.B., Simpson R.J., Ward L.D.,
 RA Knox R.B.;
 RT "Cloning and sequencing of Lol p I, the major allergenic protein of
 RT rye-grass pollen.";
 RL FEBS Lett. 279:210-215(1991).
 RN [3]
 RP SEQUENCE OF 24-53.
 RC TISSUE=Pollen;
 RX MEDLINE=86242068; PubMed=3718469;
 RA Cottam G.P., Moran D.M., Standing R.;
 RT "Physicochemical and immunochemical characterization of allergenic
 RT proteins from rye-grass (Lolium perenne) pollen prepared by a rapid
 RT Biochem. J. 234:305-310(1986).
 RL [4]
 RP SEQUENCE OF 236-263.
 RX MEDLINE=89364850; PubMed=2475768;
 RA Esch R.E., Klapper D.G.;
 RT "Isolation and characterization of a major cross-reactive grass group
 RT I allergenic determinant.";
 RL Mol. Immunol. 26:557-561(1989).
 CC -|- SUBCELLULAR LOCATION: Secreted.
 CC -|- DISEASE: CAUSES GRASS POLLEN ALLERGY.
 CC -|- SIMILARITY: BELONGS TO THE Lol p I FAMILY OF ALLERGENS.
 CC -|- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE EG45 DOMAIN.
 CC -|- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE CBD DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M57476; AAA63279.1; -;
 DR EMBL: M57476; AAA63278.1; -;
 DR PIR: A23341; A23341.
 DR PIR: B37881; B37881.
 DR PIR: S13614; S13614.
 DR HSSP: P43214; 1MHO.
 DR Interpro: IPR000882; pollen.allergen.
 DR Pfam: PF01357; pollen.allergen.1.
 DR PRINTS: PRO1225; EXPANSINFAMLY.
 DR Prodom: PD002179; pollen.allergen.1.
 DR PROSITE: PS50843; EXPANSIN_CBD; 1.
 DR PROSITE: PS50842; EXPANSIN_EG45; 1.
 KW Allergen; Glycoprotein; Signal; Multigene family.
 FT SIGNAL 1 23
 FT CHAIN 24 263 POLLEN ALLERGEN Lol p 1.
 FT DOMAIN 61 167 EXPANSIN-LIKE EG45.
 FT DOMAIN 181 262 EXPANSIN-LIKE CBD.
 FT CARBOHYD 32 32 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARIANT 68 68 N -> D.
 FT VARIANT 177 177 A -> G.
 FT VARIANT 210 210 I -> T.
 FT VARIANT 246 246 F -> V.
 FT VARIANT 28 28 P -> C (IN REF. 3).
 FT CONFLICT 31 31 P -> C (IN REF. 3).
 FT CONFLICT 48 48 W -> MW (IN REF. 3).
 FT CONFLICT 48 48 W -> MW (IN REF. 3).
 SQ SEQUENCE 263 AA; 28438 MW; 7675896F279C88C9 CAC64;
 Query Match 16.7%; Score 210; DB 1; Length 263;
 Best Local Similarity 26.9%; Pred. No. 5.5e-11;
 Matches 64; Conservative 40; Mismatches 104; Indels 30; Gaps 9;
 Qy 2 GGGWVAHATFYGGGDASGTM--GGACGCGNLYSGCGTNTALSTALPNNGLSCGACFE 59
 Db 38 GSKWLEAATFYGSGNPGCAAPDDHGGAGYKVDKPPEDMTAGCEPIPKDGLGCRACY 97
 Qy 60 IRCQNDKKWCPLPGSIVVTATNFCPPNNALPNNAGWCNPPQGHFDLSQPVFORIO--- 115
 Db 98 ICTKR-EGCSEAVYVITD-----DNEP-----IAPYHFDLSGAFAMAKKKEE 144
 Qy 116 ---YRAGIVPAVYRRVPCVRRGGIRFTIN---GHSYFNLVLTITVNGAGDVHSAWVG 167
 Db 145 QNVSAGLELQFRVYKCKYPSGKTITFEIKGSNDHYLALVYKVAAGDGNIVAVDIKEK 204
 Qy 168 SRICQAMSRNNGQNNQ--SNSYLNQSLSFVYTSDDGTIYSNNXANAGSFGQTFT 223
 Db 205 GSKWLEAATFYGSGNPGCAAPDDHGGAGYKVDKPPEDMTAGCEPIPKDGLGCRACY 261
 RESULT 3
 MPIL_PHLPR STANDARD; PRT; 263 AA.
 AC P43213;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Pollen allergen Phl p 1 precursor (Phl p I).
 GN PHLPI.
 OS Phleum pratense (Common timothy).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
 OC Poaceae; Phleum.
 OC NCBI_TaxId=15957;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pollen;
 RX MEDLINE=95015525; PubMed=7930302;
 RA Laffer S., Valente R., Vrtala S., Susani M., van Ree R., Kraft D.,
 RA Schindler O., Duchene M.;
 RT "Complementary DNA cloning of the major allergen Phl p I from timothy
 RT grass (Phleum pratense); recombinant Phl p I inhibits IgE binding to

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 11, 2002, 14:50:30 : Search time 5.8893 Seconds

(without alignments)
1499.000 Million cell updates/sec

Title: US-09-896-301-2

Perfect score: 1259

Sequence: 1 AGGWNVAHATFYGGDASG.....NNXANAGMSFGQTFGAHVR 228

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	214.5	17.0	246	1 MPCL_CYNDA	004701 cynodon dac
2	210	16.7	263	1 MPCL_LOLPR	P14946 lolium pere
3	210	16.7	263	1 MPPL_PHLPR	P43213 phleum prat
4	199	15.8	265	1 MPH1_HOLLA	P43216 holcus lana
5	199	15.8	269	1 MPAL_PHAAD	Q41260 phalaris ag
6	168.5	13.4	263	1 MPOL_ORISA	Q40638 oryza sativ
7	166	13.2	191	1 MPZ1_MAIZE	Q07154 zea mays (m
8	110.5	8.8	242	1 GUN5_TRIRE	P43317 trichoderma
9	94.5	7.5	385	1 PER_DRONE	P91686 drosophila
10	93	7.4	333	1 CAT2_MOUSE	Q91014 mus musculu
11	89	7.1	449	1 ROH2_HUMAN	P55795 homo sapien
12	88.5	7.0	405	1 YCC8_YEAST	P25367 saccharomyc
13	86.5	6.9	922	1 PMPL_CHLPP	Q92995 chlamydia p
14	86.5	6.9	1385	1 YMS1_CAEEL	P34501 caenorhabdi
15	86	6.8	181	1 YF48_MYCTU	Q10778 mycobacteri
16	85	6.8	181	1 GUN MYTED	P19469 mytilus edu
17	84.5	6.7	207	1 EGG2_SCHJA	P14585 schistosoma
18	84.5	6.7	1429	1 LIT2_CAEEL	P14585 caenorhabdi
19	84	6.7	90	1 YV77_STNY3	P73455 synechocyst
20	84	6.7	1341	1 VG37_BPT2	P07667 bacteriophag
21	84	6.7	1656	1 OMPB_RICJA	Q06653 r outer mem
22	83	6.6	443	1 PORD_PSEAE	P32722 pseudomonas
23	81	6.4	223	1 VG32_BPMD2	Q64226 mycobacteri
24	80.5	6.4	677	1 SP87_DICDI	P54643 dictyostell
25	80.5	6.4	1403	1 YGNI_YEAST	P53127 saccharomyc
26	80	6.4	392	1 Y462_TREPA	O83475 treponema p
27	79.5	6.3	1300	1 L20K_RICRI	P14914 rickettsia
28	79.5	6.3	1654	1 OMPB_RICRI	Q53047 r outer mem
29	79	6.3	112	1 SECG_HAEIN	P44713 haemophilus
30	79	6.3	212	1 EGG1_SCHJA	P19470 schistosoma
31	79	6.3	449	1 ROH1_HUMAN	P31943 homo sapien
32	79	6.3	465	1 GRP2_PHAUV	P10496 phaseolus v
33	79	6.3	751	1 VP4_ROTGI	P15155 rotavirus (

34	78.5	6.2	391	1 PER_DROIN	P91613 drosophila
35	78.5	6.2	564	1 MERA_SHIFL	P08332 shigella fl
36	78.5	6.2	666	1 MUR2_ENTHR	P39046 enterococcu
37	78.5	6.2	1645	1 OMPB_RICRY	P96989 r outer mem
38	78	6.2	354	1 ALC_NEUCR	P18407 neurospora
39	78	6.2	378	1 ELYA_BACSP	P20724 bacillus sp
40	78	6.2	481	1 LORI_MOUSE	P18165 mus musculu
41	78	6.2	645	1 K22E_HUMAN	P35908 homo sapien
42	78	6.2	1246	1 YWV2_CAEEL	P34504 caenorhabdi
43	78	6.2	1964	1 NMC4_MOUSE	P31695 mus musculu
44	77.5	6.2	712	1 CPGT_BAC53	P09121 bacillus sp
45	77	6.1	571	1 FLAL_CAME	P56963 campylobact

ALIGNMENTS

RESULT 1	ID	MPCL_CYNDA	STANDARD:	PRT:	246 AA.
AC	004701:				
DT	16-OCT-2001 (rel. 40, Created)				
DT	16-OCT-2001 (rel. 40, Last sequence update)				
DE	01-MAR-2002 (rel. 41, Last annotation update)				
DE	Major pollen allergen Cyn d 1.				
GN	CYNOL				
OS	Cynodon dactylon (Bermuda grass).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;				
OC	Chloridoideae; Cynodonteae; Cynodon.				
OX	NCBI_TaxID=28909;				
RN	[1]				
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.				
RC	TISSUE=Pollen.				
RX	MEDLINE=96347957; PubMed=8757211;				
RA	Smith P.M., Supphloglu C., Griffith I.J., Theriault K., Knox R.B.,				
RA	Singh M.B.;				
RT	*Cloning and expression in yeast Pichia pastoris of a biologically				
RT	active form of Cyn d 1, the major allergen of Bermuda grass pollen.;				
RL	J. Allergy Clin. Immunol. 98:331-343(1996).				
CC	-1- SUBCELLULAR LOCATION: Secreted.				
CC	-1- DISEASE: CAUSES GRASS POLLEN ALLERGY.				
CC	-1- SIMILARITY: BELONGS TO THE LOL P I FAMILY OF ALLERGENS.				
CC	-1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE EG45 DOMAIN.				
CC	-1- SIMILARITY: CONTAINS 1 EXPANSIN-LIKE CBD DOMAIN.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/				
CC	or send an email to license@sib-sib.ch).				
DR	EMBL: S83343; AAB50734.2; -				
DR	HSSP: P43214; 1WHO.				
DR	InterPro: IPR000882; Pollen_allergen.				
DR	Pfam: PF01357; Pollen_allergen.1.				
DR	PRINTS: PR01225; EXPANSINFAMILY.				
DR	ProDom: PD002179; Pollen_allergen.1.				
DR	PROSITE: PS50843; EXPANSIN_CBD; 1.				
DR	PROSITE: PS50842; EXPANSIN_EG45; 1.				
KW	Allergen.				
FT	DOMAIN 39 145 EXPANSIN-LIKE EG45.				
FT	DOMAIN 159 240 EXPANSIN-LIKE CBD.				
FT	CARBOHYD 9 N-LINKED (GLCNAC. . .) (POTENTIAL).				
SQ	SEQUENCE 246 AA; 26888 MW; 43D8442DBA588322 CRC64;				
QY	2 GGWNVAHATFYGG--GDASGTMGACGYNLYSGYGTNTALSTALFNNGLSGACF 58				

Query Match Score 214.5; DB 1; Length 246;

Best local similarity 25.9%; Pred. No. 21e-11;

Matches 62; Conservative 44; Mismatches 102; Indels 31; Gaps 8;

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
 OX NCBI_TaxID=3352;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=HYPOCOTYL;
 RA Hutchison K.W., Singer P.B., Diaz-Sala C., Greenwood M.S.;
 RT "Expansins are conserved in conifers and expressed in response to
 exogenous auxin."
 RL Submitted (Jul-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U64890; AAA0634.1; ..
 DR InterPro: IPR000882; Pollen_allergen.
 DR Pfam: PF01357; Pollen_allergen; 1.
 DR PRINTS: PR01225; EXPANSNFAMLY.
 DR ProDom: PD002179; Pollen_allergen; 1.
 FT NON_TER 1
 SQ SEQUENCE 232 AA; 24734 MW; 1340AF483DE0741F CRC64;

Query Match 82.8%; Score 1042; DB 10; Length 232;
 Best Local Similarity 79.1%; Pred. No. 5.2e-87;
 Matches 178; Conservative 25; Mismatches 22; Indels 0; Gaps 0;

QY 3 GGVNNAHATFYGGDASGTMGACGYGNLYSGYGTNTALSTALFNNGLSGACFEIRC 62
 ||| : ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| : |||
 Db 8 GGWESAHAATFYGGSDASGTMGACGYGNLYSGYGTNTALSTALFNNDGLSCGACYEMRC 67

QY 63 QNDGKWCCLPGSIYVATNFPCPPNNALPNNAGWCNPPQOHFDLSOPVFORIAQYRAGIYP 122
 : : ||||| : ||||| ||||| ||||| ||||| ||||| ||||| : ||| : ||| |||||
 Db 68 NDDPQMCLEFGIVYVATNFCPPNNALPNDNGWCNPPLOHFDPAEPALFLIAKYRGIVP 127

QY 123 VAVRRVPCVVRGIRPTINGHSYFNLYLITNVGAGDVHSAMVKGSRGTGQAMSRNMGON 182
 : ||||| : ||||| ||||| ||||| ||||| ||||| ||||| : ||| : ||| |||||
 Db 128 ILYTRVPCLEKRGIRFTVNGHSYFNLYLITNVGAGDVHAVSIKSGMSGMGQPMNRNMGON 187

QY 183 WQSNSTLNGQSUSFKYTTSDGQTIYSNNXANAGWSFGQTFTGAHV 227
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| : ||| : ||| |||||
 Db 188 WQSNSTLNGQSLSFQYTTSDGRTIVSNVAPSNWQFGQTFEGSQV 232

Search completed: October 11, 2002, 14:59:39
 Job time : 21.6871 secs

DE EXPANSIN (FRAGMENT).
 OS Pinus taeda (loblolly pine).
 OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
 OC Spermatophyta: Coniferopsida: Coniferales: Pinaceae: Pinus.
 NCBI_TaxID=3352;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=HYPOCOTYL;
 RA Hutchison K.W., Singer P.B., Diaz-Sala C., Greenwood M.S.;
 RT "Expansins are conserved in conifers and expressed in response to
 exogenous auxin." to the EMBL/GenBank/DBJ databases.
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U64891; AAB40635.1; -
 DR InterPro: IPR000882; Pollen_allergen.
 DR Pfam: PF01357; Pollen_allergen.1.
 DR PRINTS: PR01225; EXPANSINFA1.
 DR ProDom: PD002179; Pollen_allergen.1.
 PT NON_TER
 SQ SEQUENCE 232 AA: 24745 MW: 1340805730DE907CF CRC64;

Query Match 83.2%; Score 1047; DB 10; Length 232;
 Best Local Similarity 79.1%; Pred. No. 18e-87;
 Matches 178; Conservative 26; Mismatches 21; Indels 0; Gaps 0;

QY 3 GGVNNAHATFYGGDASCTMGACGYNLVSQGYGTNTAALSTALFNNGLSGACFEIRC 62
 ||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||
 DB 8 GGVNNAHATFYGGDASCTMGACGYNLVSQGYGTNTAALSTALFNNGLSGACFEIRC 67
 QY 63 QNDGKWCLEPGSTIVTATNFCPPNNALPNNAGGMCNPPQHFDSQVFORIAOYRAGIYV 122
 : :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||
 DB 68 NDDPKWCLPGSTIVTATNFCPPNNALPNDNGMCNPPQHFDEMAEPFLKIAVYRAGIYV 127
 QY 123 VAYRRVPCVRRGIGFTTNGHSYFNLVLTNNVGAGDVHSAMVKSRTGQAMSRNMGON 182
 : :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||
 DB 128 ILTTRVPCRLKRGIRFTYNGHSYFNLVLTNNVGAGDVHSAVSKSRSGMQPSRNMGN 187
 QY 183 WQNSNYLNGQSLSFKVYTTSDGQTIIVSNXNAGMSFGQTFEGAHV 227
 ||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||
 DB 188 WQNSNYLNGQSLSFQVTTSDGRTIVSNVAPSNMFGQTFEGSQV 232

RESULT 13
 Q9SWY1 PRELIMINARY; PRT; 253 AA.
 AC Q9SWY1;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE EXPANSIN.
 OS Pinus taeda (loblolly pine).
 OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
 OC Spermatophyta: Coniferopsida: Coniferales: Pinaceae: Pinus.
 NCBI_TaxID=3352;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=HYPOCOTYL;
 RA MEDLINE=9939318; PubMed=10398718;
 RA Hutchison K.W., Singer P.B., McInnis S., Diaz-Sala C., Greenwood M.S.;
 RT "Expansins are conserved in conifers and expressed in hypocotyls in
 response to exogenous auxin.";
 RL Plant Physiol. 120:827-832(1999).
 DR EMBL: AF085330; AAD47901.1; -
 DR InterPro: IPR000882; Pollen_allergen.
 DR Pfam: PF01357; Pollen_allergen.1.
 DR PRINTS: PR01225; EXPANSINFA1.
 DR ProDom: PD002179; Pollen_allergen.1.
 SQ SEQUENCE 253 AA: 27150 MW: 267581569BD75D4 CRC64;

Query Match 83.2%; Score 1047; DB 10; Length 253;
 Best Local Similarity 79.1%; Pred. No. 2e-87;
 Matches 178; Conservative 26; Mismatches 21; Indels 0; Gaps 0;

QY 3 GGVNNAHATFYGGDASCTMGACGYNLVSQGYGTNTAALSTALFNNGLSGACFEIRC 62

DB 29 GGVNNAHATFYGGDASCTMGACGYNLVSQGYGTNTAALSTALFNNGLSGACFEIRC 88
 ||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||
 QY 63 QNDGKWCLEPGSTIVTATNFCPPNNALPNNAGGMCNPPQHFDSQVFORIAOYRAGIYV 122
 : :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||
 DB 89 NDDPKWCLPGSTIVTATNFCPPNNALPNDNGMCNPPQHFDEMAEPFLKIAVYRAGIYV 148
 QY 123 VAYRRVPCVRRGIGFTTNGHSYFNLVLTNNVGAGDVHSAMVKSRTGQAMSRNMGON 182
 : :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||
 DB 149 ILTTRVPCRLKRGIRFTYNGHSYFNLVLTNNVGAGDVHSAVSKSRSGMQPSRNMGN 208
 QY 183 WQNSNYLNGQSLSFKVYTTSDGQTIIVSNXNAGMSFGQTFEGAHV 227
 ||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||
 DB 209 WQNSNYLNGQSLSFQVTTSDGRTIVSNVAPSNMFGQTFEGSQV 253

RESULT 14
 Q9LUB2 PRELIMINARY; PRT; 245 AA.
 AC Q9LUB2;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE EXPANSIN 2.
 OS Zinnia elegans.
 OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
 OC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots:
 OC Asterales: euasterids II: Asterales: Asteraceae: Asteroideae:
 OC Heliantheae: Zinnia.
 NCBI_TaxID=34245;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20317189; PubMed=10859177;
 RA Im K.H., Cosgrove D.J., Jones A.M.;
 RT "Subcellular localization of expansin mRNA in xylem cells.";
 RL Plant Physiol. 123:463-470(2000).
 DR EMBL: AF230332; AAF35901.1; -
 DR InterPro: IPR000882; Pollen_allergen.
 DR Pfam: PF01357; Pollen_allergen.1.
 DR PRINTS: PR01225; EXPANSINFA1.
 DR ProDom: PD002179; Pollen_allergen.1.
 SQ SEQUENCE 245 AA: 26201 MW: F66E2611AC27E360 CRC64;

Query Match 83.1%; Score 1046; DB 10; Length 245;
 Best Local Similarity 81.1%; Pred. No. 2.4e-87;
 Matches 180; Conservative 22; Mismatches 20; Indels 0; Gaps 0;

QY 3 GGVNNAHATFYGGDASCTMGACGYNLVSQGYGTNTAALSTALFNNGLSGACFEIRC 62
 ||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||
 DB 21 GGVNNAHATFYGGDASCTMGACGYNLVSQGYGTNTAALSTALFNNGLSGACFEIRC 80
 QY 63 QNDGKWCLEPGSTIVTATNFCPPNNALPNNAGGMCNPPQHFDSQVFORIAOYRAGIYV 122
 : :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||
 DB 81 NDDPKWCLPGSTIVTATNFCPPNNALPNDNGMCNPPQHFDEMAEPFLKIAVYRAGIYV 140
 QY 123 VAYRRVPCVRRGIGFTTNGHSYFNLVLTNNVGAGDVHSAMVKSRTGQAMSRNMGON 182
 : :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||
 DB 141 ISFORVPCRLKRGIRFTYNGHSYFNLVLTNNVGAGDVHSAVSKSRSGMQPSRNMGN 200
 QY 183 WQNSNYLNGQSLSFKVYTTSDGQTIIVSNXNAGMSFGQTFEGAHV 224
 ||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||
 DB 201 WQNSNYLNGQSLSFQVTTSDGRTIVSNVAPSNMFGQTFEGSQV 242

RESULT 15
 P93492 PRELIMINARY; PRT; 232 AA.
 AC P93492;
 DT 01-MAY-1997 (TREMBLrel. 03, Created)
 DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE EXPANSIN (FRAGMENT).
 OS Pinus taeda (loblolly pine).

DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
 DE ALPHA-EXPANSIN 3.
 OS Triphysaria versicolor.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; eusterids I; Lamiales; Orobanchaceae; Triphysaria.
 NCBI_TaxID=64093;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=ROOTS TREATED WITH DMBQ;
 RX MEDLINE=21186069; PubMed=11290422;
 RA Wrobel R.L., Yoder J.I.;
 RT "Differential RNA expression of alpha-expansin gene family members in
 the parasitic angiosperm *Triphysaria versicolor* (Scrophulariaceae).";
 RL Gene 266:85-93(2001).
 DR EMBL: AF230276; AAF32409.1; -
 DR InterPro: IPR000882; Pollen_allergen.
 DR PRINTS: PRO1225; EXPANSNPFAMLY.
 DR Prodom: PD002179; Pollen_allergen.1.
 SQ SEQUENCE 247 AA; 26456 MW; 0B2AC8F357CCE7B5 CRC64;
 Query Match 86.6%; Score 1090; DB 10; Length 247;
 Best Local Similarity 85.0%; Pred. No. 2,4e-91;
 Matches 192; Conservative 14; Mismatches 20; Indels 0; Gaps 0;
 QY 3 GGVNNAHATFFGGDASGTMGACGCGNLYSGYGTNTAALSTALFNNGLSGACPEIRG 62
 DB 22 GGVNNAHATFFGGDASGTMGACGCGNLYSGYGTNTAALSTALFNNGLSGACPEIRG 81
 QY 63 QNGKMCLEPGSIYVTAATNCPNNALPNNAGWCNPPQOHFDLSQVFORIAQYRAGIVP 122
 DB 82 VNDGKMCLEPGSIYVTAATNCPNNALPNNAGWCNPPHFDLAQVFOHIAQYRAGIVP 141
 QY 123 VAYRVPVCVRGGIRPTINGHSYFNLYLTNVGAGADVHSAMVKSRTGQMSRWGON 182
 DB 142 VAYRVPVCVRGGIRPTINGHSYFNLYLTNVGAGADVHSAMVKSRTGQMSRWGON 201
 QY 183 WQSNSTLNGOSLSFKVTTSDGQIVSNXNANAGMSFGQFTGAHVR 228
 DB 202 WQSNSTLNGOSLSFKVTTSDGQIVSNXNANAGMSFGQFTGAHVR 247
 RESULT 10
 Q9LBI PRELIMINARY: PRT; 242 AA.
 AC Q9LBI;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)
 DE EXPANSIN 3.
 OS Zinnia elegans.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; eusterids II; Asterales; Asteraceae; Asteroideae;
 OC Heliantheae; Zinnia.
 NCBI_TaxID=34245;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20317189; PubMed=10859177;
 RA Im R.H., Cosgrove D.J., Jones A.M.;
 RT "Subcellular localization of expansin mRNA in xylem cells.";
 RL Plant Physiol. 123:463-470(2000).
 DR EMBL: AF230333; AAF35902.1; -
 DR InterPro: IPR000882; Pollen_allergen.
 DR Pfam: PF01357; Pollen_allergen.1.
 DR PRINTS: PD001225; EXPANSNPFAMLY.
 DR Prodom: PD002179; Pollen_allergen.1.
 SQ SEQUENCE 242 AA; 26157 MW; 168A01724FCF5B58 CRC64;
 Query Match 83.5%; Score 1051.5; DB 10; Length 242;
 Best Local Similarity 83.5%; Pred. No. 7.5e-88;
 Matches 187; Conservative 15; Mismatches 17; Indels 5; Gaps 1;

QY 2 GGVNNAHATFFGGDASGTMGACGCGNLYSGYGTNTAALSTALFNNGLSGACPEIR 61
 DB 21 GGVNNAHATFFGGDASGTMGACGCGNLYSGYGTNTAALSTALFNNGLSGACPEIR 80
 QY 62 CONDKMCLEPGSIYVTAATNCPNNALPNNAGWCNPPQOHFDLSQVFORIAQYRAGIV 121
 DB 81 CVNDRQWCLPATIVTATNFCP-----PNSNGWCDDPRQHFHDSQPIFGHIAQKAVV 135
 QY 122 PVAYRVPVCVRGGIRPTINGHSYFNLYLTNVGAGADVHSAMVKSRTGQMSRWGON 181
 DB 136 PVAYRVPVCVRGGIRPTINGHSYFNLYLTNVGAGADVHSAMVKSRTGQMSRWGON 195
 QY 182 NMQSNSTLNGOSLSFKVTTSDGQIVSNXNANAGMSFGQFTGA 225
 DB 196 NMQSNSTLNGOSLSFKVTTSDGQIVSNXNANAGMSFGQFTGA 239
 RESULT 11
 Q9FVG9 PRELIMINARY: PRT; 250 AA.
 AC Q9FVG9;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
 DE EXPANSIN.
 OS Lycopersicon esculentum (Tomato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; eusterids I; Solanales; Solanaceae; Solanum.
 NCBI_TaxID=4081;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20532800; PubMed=11080302;
 RA Chen F., Bradford K.J.;
 RT "Expression of an Expansin Is Associated with Endosperm Weakening
 during Tomato Seed Germination.";
 RL Plant Physiol. 124:1265-1274(2000).
 DR EMBL: AF184233; AAG32921.1; -
 DR InterPro: IPR000882; Pollen_allergen.
 DR PRINTS: PRO1225; EXPANSNPFAMLY.
 DR Prodom: PD002179; Pollen_allergen.1.
 SQ SEQUENCE 250 AA; 26821 MW; ADD1508763680CFD CRC64;
 Query Match 83.3%; Score 1049; DB 10; Length 250;
 Best Local Similarity 81.5%; Pred. No. 1.3e-87;
 Matches 181; Conservative 22; Mismatches 19; Indels 0; Gaps 0;
 QY 4 GGVNNAHATFFGGDASGTMGACGCGNLYSGYGTNTAALSTALFNNGLSGACPEIRG 63
 DB 26 GGVNNAHATFFGGDASGTMGACGCGNLYSGYGTNTAALSTALFNNGLSGACPEIRG 85
 QY 64 NDGKMCLEPGSIYVTAATNCPNNALPNNAGWCNPPQOHFDLSQVFORIAQYRAGIVP 123
 DB 86 NDGKMCLEPGSIYVTAATNCPNNALPNNAGWCNPPHFDLAQVFOHIAQYRAGIVP 145
 QY 124 AYRVPVCVRGGIRPTINGHSYFNLYLTNVGAGADVHSAMVKSRTGQMSRWGON 183
 DB 146 AYRVPVCVRGGIRPTINGHSYFNLYLTNVGAGADVHSAMVKSRTGQMSRWGON 205
 QY 184 QSNSTLNGOSLSFKVTTSDGQIVSNXNANAGMSFGQFTGA 225
 DB 206 QSNSTLNGOSLSFKVTTSDGQIVSNXNANAGMSFGQFTGA 247
 RESULT 12
 P93493 PRELIMINARY: PRT; 232 AA.
 AC P93493;
 DT 01-MAY-1997 (TREMBLrel. 03, Created)
 DT 01-MAY-1997 (TREMBLrel. 03, last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)

RA Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhagen G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA Selinger S.L., Fraser C.M., Venter J.C.,
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
thaliana".
RT Nature 402:761-768(1999).
RN
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC004138; AAC32927.1; -;
DR InterPro: IPR000882; Pollen_allergen.
DR Pfam: PF01357; Pollen_allergen; 1.
DR PRINTS: PR01225; EXPANSNFAMLY.
DR ProDom: PD002179; Pollen_allergen; 1.
SQ SEQUENCE 248 AA: 26482 MW: 29E3199269B71271 CRC64;

Query Match 88.6%; Score 1115; DB 10; Length 248;
Best Local Similarity 87.1%; Pred. No. 1.3e-93;
Matches 196; Conservative 13; Mismatches 16; Indels 0; Gaps 0;

QY 4 GGWNAHATFFGGDASGTMGACGYGNTLSOGYGTNTALSTALFNNGLSGACFEIRCO 63
DB 24 GGWNAHATFFGGDASGTMGACGYGNTLSOGYGTNTALSTALFNNGLSGACFEIRCO 83
QY 64 NDGKWLPGSIVTATNFCPPNNALPNNAGGWCNPPQOHFDLSQPVFORIAQYRAGIVP 123
DB 84 SDGKWLPGSIVTATNFCPPNNALPNNAGGWCNPPQOHFDLSQPVFORIAQYRAGIVP 143
QY 124 AYRRVPCVRKRGIRFTTNGHSYFNLVLTNVGAGDVHSAVWGSRGQMSRNMGON 183
DB 144 SYRRVPCVRKRGIRFTTNGHSYFNLVLTNVGAGDVHSAVWGSRGQMSRNMGON 203
QY 184 OSNLSYNGOSLSFKVTTSDGOTIVSNXNAGSFGOTFTGANVR 228
DB 204 OSNLSYNGOSLSFKVTTSDGOTIVSNXNAGSFGOTFTGANVR 248

RESULT 7
ID 09M515 PRELIMINARY; PRT; 249 AA.
AC 09M515;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ALPHA-EXPANSIN 1.
OS Tripharyxalis versicolor.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; eusterids I; Lamiales; Orobanchaceae; Tripharyxalis.
OX NCBI_TaxID=64093;
RN
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=ROOTS TREATED WITH DMBQ;
RA MEDLINE-21186069; PubMed-11290422;
RA Wrobel R.L., Yoder J.I.;
RT "Differential RNA expression of alpha-expansin gene family members in
the parasitic angiosperm Tripharyxalis versicolor (Scrophulariaceae).";
RL Gene 266:85-93(2001).
DR EMBL: AF230278; AF232411.1; -;
DR InterPro: IPR000882; Pollen_allergen.
DR PRINTS: PR01225; EXPANSNFAMLY.
DR ProDom: PD002179; Pollen_allergen; 1.
SQ SEQUENCE 249 AA: 26116 MW: DC976FEE376794FF CRC64;

Query Match 87.9%; Score 1106.5; DB 10; Length 249;
Best Local Similarity 87.5%; Pred. No. 7.7e-93;
Matches 196; Conservative 17; Mismatches 10; Indels 1; Gaps 1;
QY 3 GGWNAHATFFGGDASGTMGACGYGNTLSOGYGTNTALSTALFNNGLSGACFEIRCO 62
DB 24 GGWNAHATFFGGDASGTMGACGYGNTLSOGYGTNTALSTALFNNGLSGACFEIRCO 83

DB 24 GGWNAHATFFGGDASGTMGACGYGNTLSOGYGTNTALSTALFNNGLSGACFEIRCO 83
QY 63 ONDGKWLPGSIVTATNFCPPNNALPNNAGGWCNPPQOHFDLSQPVFORIAQYRAGIVP 122
DB 84 VNDGKWLPGSIVTATNFCPPNNALPNNAGGWCNPPQOHFDLSQPVFORIAQYRAGIVP 143
QY 123 VYRRVPCVRKRGIRFTTNGHSYFNLVLTNVGAGDVHSAVWGSRGQMSRNMGON 182
DB 144 VYRRVPCVRKRGIRFTTNGHSYFNLVLTNVGAGDVHSAVWGSRGQMSRNMGON 203
QY 183 KOSNLSYNGOSLSFKVTTSDGOTIVSNXNAGSFGOTFTGANVR 225
DB 204 KOSNLSYNGOSLSFKVTTSDGOTIVSNXNAGSFGOTFTGANVR 247

RESULT 8
ID 09P36 PRELIMINARY; PRT; 249 AA.
AC 09P36;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ALPHA-EXPANSIN PRECURSOR.
GN NT-EXP4.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; eusterids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. BY2;
RA MEDLINE-99026292; PubMed-9808735;
RA Link B.M., Cosgrove D.J.;
RT "Acid-growth response and alpha-expansins in suspension cultures of
bright yellow 2 tobacco".
RL Plant Physiol. 118:907-916(1998).
DR EMBL: AF049353; AAC96080.1; -;
DR InterPro: IPR000882; Pollen_allergen.
DR Pfam: PF01357; Pollen_allergen; 1.
DR PRINTS: PR01225; EXPANSNFAMLY.
DR ProDom: PD002179; Pollen_allergen; 1.
KW Signal.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 249 ALPHA-EXPANSIN.
SQ SEQUENCE 249 AA: 26597 MW: 98F1D583B5E846CA CRC64;

Query Match 86.7%; Score 1092; DB 10; Length 249;
Best Local Similarity 85.3%; Pred. No. 1.6e-91;
Matches 191; Conservative 16; Mismatches 17; Indels 0; Gaps 0;
QY 2 GGWNAHATFFGGDASGTMGACGYGNTLSOGYGTNTALSTALFNNGLSGACFEIR 61
DB 23 GGWNAHATFFGGDASGTMGACGYGNTLSOGYGTNTALSTALFNNGLSGACFEIR 82
QY 62 CONDGKWLPGSIVTATNFCPPNNALPNNAGGWCNPPQOHFDLSQPVFORIAQYRAGIV 121
DB 83 VNDGKWLPGSIVTATNFCPPNNALPNNAGGWCNPPQOHFDLSQPVFORIAQYRAGIV 142
QY 122 VYRRVPCVRKRGIRFTTNGHSYFNLVLTNVGAGDVHSAVWGSRGQMSRNMGON 181
DB 143 VYRRVPCVRKRGIRFTTNGHSYFNLVLTNVGAGDVHSAVWGSRGQMSRNMGON 202
QY 182 NMOSNLSYNGOSLSFKVTTSDGOTIVSNXNAGSFGOTFTGANVR 225
DB 203 NMOSNLSYNGOSLSFKVTTSDGOTIVSNXNAGSFGOTFTGANVR 246

RESULT 9
ID 09M517 PRELIMINARY; PRT; 247 AA.
AC 09M517;
DT 01-OCT-2000 (TREMBLrel. 15, Created)

RC STRAIN=CV, COLUMBIA;
RA Duracko D.M., Sheherban T.Y., Cosgrove D.J.;
RT *Atxaplo is expressed in the trichomes, petioles, midribs, and pedicel
RT abscission zone."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC006535; AAF87031.1; -
DR EMBL; AF229437; AAF61712.1; -
DR EMBL; AF229431; AAF61713.1; -
DR InterPro: IPR000882; Pollen_allergen.
DR PRINTS; PRO1225; EXPANSINFAMILY.
DR ProDom; PD002179; Pollen_allergen; 1.
SQ SEQUENCE 249 AA; 26428 MW; 63D014410DABBA1 CRC64;

Query Match 90.5%; Score 1139; DB 10; Length 249;
Best Local Similarity 89.0%; Pred. No. 8.6e-96;
Matches 202; Conservative 12; Mismatches 13; Indels 0; Gaps 0;

QY 2 GGGWVNAHATFYGGDASGTMGACGYGNLYSGYGTNTAALSTALFNNGLSGACFEIR 61
DB 23 GGGWVNAHATFYGGDASGTMGACGYGNLYSGYGTNTAALSTALFNNGLSGACFEIR 82
QY 62 CONDKWKLPGSIVTATNFCPPNNALPNNAGWCNPPQOHFDLSQPVFORIAQYRAGIV 121
DB 83 CENDGKWCLEPGSIVTATNFCPPNNALANNNGWCNPPLEHFDLAQPVFORIAQYRAGIV 142
QY 122 PVAYRRVPCVRRGIRFTINGHSYFNLVLTITNGAGDVHSANVKSRTGQWMSRNMQ 181
DB 143 PVAYRRVPCVRRGIRFTINGHSYFNLVLTITNGAGDVHSAAIKSRVWQMSRNMQ 202
QY 182 NMQSNSTYLNQGLSFKVTTSDGQTTIVSNXNAGWCFQGTFTGAHR 228
DB 203 NMQSNSTYLNQGLSFKVTTSDGRTVVSFNAAPAGWCFQGTFTGAHR 249

RESULT 4

Q9FUM2 PRELIMINARY; PRT; 252 AA.
AC Q9FUM2;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE EXPANSIN 2.
GN EXP2.
OS Prunus avium (Cherry).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.
OX NCBI_TaxID=42229;
RN 11
RP SEQUENCE FROM N.A.
RA Wu Z., Wiersma P.A.;
RT "Differential Expression of Expansin Genes Isolated from Sweet Cherry
RT (Prunus avium L.) During Fruit Ripening."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF297522; AAG13983.1; -
DR InterPro: IPR000882; Pollen_allergen.
DR Pfam; PF01357; Pollen_allergen; 1.
DR PRINTS; PRO1225; EXPANSINFAMILY.
DR ProDom; PD002179; Pollen_allergen; 1.
SQ SEQUENCE 252 AA; 26766 MW; 5C62EBF83E5A138E CRC64;

Query Match 90.3%; Score 1137; DB 10; Length 252;
Best Local Similarity 89.0%; Pred. No. 1.3e-95;
Matches 202; Conservative 11; Mismatches 14; Indels 0; Gaps 0;

QY 2 GGGWVNAHATFYGGDASGTMGACGYGNLYSGYGTNTAALSTALFNNGLSGACFEIR 61
DB 26 GGGWVNAHATFYGGDASGTMGACGYGNLYSGYGTNTAALSTALFNNGLSGACFEIR 85
QY 62 CONDKWKLPGSIVTATNFCPPNNALPNNAGWCNPPQOHFDLSQPVFORIAQYRAGIV 121
DB 86 CVDPRKWCLEPGSIVTATNFCPPNNALPNNAGWCNPPQOHFDLSQPVFORIAQYRAGIV 145

QY 122 PVAYRRVPCVRRGIRFTINGHSYFNLVLTITNGAGDVHSANVKSRTGQWMSRNMQ 181
DB 146 PVAYRRVPCVRRGIRFTINGHSYFNLVLTITNGAGDVHSANVKSRTGQWMSRNMQ 205
QY 182 NMQSNSTYLNQGLSFKVTTSDGQTTIVSNXNAGWCFQGTFTGAHR 228
DB 206 NMQSNSTYLNQGLSFKVTTSDGRTVVSFNAAPAGWCFQGTFTGAHR 252

RESULT 5

Q93XPL PRELIMINARY; PRT; 249 AA.
AC Q93XPL;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE EXPANSIN.
GN EXP3.
OS Prunus cerasus.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.
OX NCBI_TaxID=140311;
RN 11
RP SEQUENCE FROM N.A.
RA TISSUE=FRUIT ABSCISSION ZONE;
RA Yoo S.-D., van Nocker S.;
RL Submitted (Feb-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF350938; AAK4847.1; -
SQ SEQUENCE 249 AA; 26397 MW; C65556B038AB99D0 CRC64;

Query Match 89.0%; Score 1120; DB 10; Length 249;
Best Local Similarity 87.5%; Pred. No. 4.6e-94;
Matches 196; Conservative 16; Mismatches 12; Indels 0; Gaps 0;

QY 1 AGGGWVNAHATFYGGDASGTMGACGYGNLYSGYGTNTAALSTALFNNGLSGACFEI 60
DB 23 SGGWVNAHATFYGGDASGTMGACGYGNLYSGYGTNTAALSTALFNNGLSGACFEI 82
QY 61 RCONDKWKLPGSIVTATNFCPPNNALPNNAGWCNPPQOHFDLSQPVFORIAQYRAGI 121
DB 83 RCASDPKWLPGSIVTATNFCPPNNALPNNAGWCNPPQOHFDLAQPVFORIAQYRAGI 142
QY 121 PVAYRRVPCVRRGIRFTINGHSYFNLVLTITNGAGDVHSANVKSRTGQWMSRNMQ 180
DB 143 VVAYRRVPCVRRGIRFTINGHSYFNLVLTITNGAGDVHSAAIKSRVWQMSRNMQ 202
QY 181 NMQSNSTYLNQGLSFKVTTSDGQTTIVSNXNAGWCFQGTFTGAHR 224
DB 203 NMQSNSTYLNQGLSFKVTTSDGRTVVSFNAAPAGWCFQGTFTGAHR 246

RESULT 6

Q80622 PRELIMINARY; PRT; 248 AA.
AC Q80622;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PUTATIVE EXPANSIN.
GN AT2G03090.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN 11
RP SEQUENCE FROM N.A.
RA STRAIN=CV, COLUMBIA;
RA MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,

QY 61 RCNDGKMKCLPGSIIVYATNCPNNALPNNAGMCNPPQHFDSLQPVFORIAQYRAGI 120
 DB 70 RCNDGKMKCLPGSIIVYATNCPNNALPNNAGMCNPPQHFDSLQPVFORIAQYRAGI 129
 QY 121 VPAVARRVPCVRCGRIFFTINGHSYFNLVLTNNVGAGDVHSAMVKGSRFGQMSRNMKG 180
 DB 130 VPAVARRVPCVRCGRIFFTINGHSYFNLVLTNNVGAGDVHSAMVKGSRFGQMSRNMKG 189
 QY 181 QNMOSNSTYLNQSLSFVKYTTSDGQITVSNXNAGMSEFGQTFPGAHR 228
 DB 190 QNMOSNSTYLNQSLSFVKYTTSDGQITVSNXNAGMSEFGQTFPGAHR 237
 RESULT 2
 Q9C554 PRELIMINARY: PRT: 250 AA.
 AC Q9C554:
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE PUTATIVE EXPANSIN PROTEIN (AT-EXP1).
 GN F10D13_18.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
 OC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae:
 OC Eurosid II: Brassicales: Brassicaceae: Arabidopsids.
 OX NCBI_TaxID=3702;
 RP SEQUENCE FROM N.A.
 RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
 RA Goldsmith A.D., Lee J.M., Quach H.L., Tortum M., Yu G., Bowser L.,
 RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
 RA Kamuya A., Karlín-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
 RA Miranda M., Narusaka M., Nguyen N., Palm C.J., Sakurai T., Satou M.,
 RA Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
 RA Theologis A.:
 RT "Full length cDNA of gene F10D13.18 (GI:12597783)."
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA MEDLINE=21016719; PubMed=11130712;
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
 RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
 RA Buchler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
 RA Dunn P., Elgu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
 RA Gull J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
 RA Kim C.J., Koo H.L., Kremetsekis I., Kurtz D.B., Kwan A., Lam B.,
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marzilli A.,
 RA Miltner J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
 RA Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
 RA Sun H., Tallon L.J., Tambunga G., Tortum M.J., Town C.D.,
 RA Uterback T., Van Aken S., Vaysberg M., Vysotskaya V.S., Walker M.,
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.:
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 thaliana."
 RL Nature 408:816-820(2000).
 DR EMBL: AF360291; AAR6001.1; -;
 DR EMBL: AC073178; AAC60095.1; -;
 DR Interpro: IPR000882; Pollen_allergen.
 DR PRINTS: PRO1225; EXPANSINFAMLY.
 DR PRODOM: PD002179; Pollen_allergen; 1.
 SO SEQUENCE 250 AA; ID95EBA24FCFET5 CRC64;
 Query Match 99.0%; Score 1246; DB 10; Length 250;
 Best Local Similarity 98.7%; Pred. NO.1.6e-105;
 Matches 225; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 OY 1 AGGGMVNAHATFYGGDASGTMGACGYNLVSQGYGTNTAALSTALFNGLSCGACFEI 60

DB 23 AGGGMVNAHATFYGGDASGTMGACGYNLVSQGYGTNTAALSTALFNGLSCGACFEI 82
 QY 61 RCNDGKMKCLPGSIIVYATNCPNNALPNNAGMCNPPQHFDSLQPVFORIAQYRAGI 120
 DB 83 RCNDGKMKCLPGSIIVYATNCPNNALPNNAGMCNPPQHFDSLQPVFORIAQYRAGI 142
 QY 121 VPAVARRVPCVRCGRIFFTINGHSYFNLVLTNNVGAGDVHSAMVKGSRFGQMSRNMKG 180
 DB 143 VPAVARRVPCVRCGRIFFTINGHSYFNLVLTNNVGAGDVHSAMVKGSRFGQMSRNMKG 202
 QY 181 QNMOSNSTYLNQSLSFVKYTTSDGQITVSNXNAGMSEFGQTFPGAHR 228
 DB 203 QNMOSNSTYLNQSLSFVKYTTSDGQITVSNXNAGMSEFGQTFPGAHR 250
 RESULT 3
 Q9LDR9 PRELIMINARY: PRT: 249 AA.
 AC Q9LDR9:
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE T24P13.15 (EXPANSIN 10).
 GN EXP10.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
 OC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae:
 OC Eurosid II: Brassicales: Brassicaceae: Arabidopsids.
 OX NCBI_TaxID=3702;
 RP SEQUENCE FROM N.A.
 RA Johnson-Hopson C., Dunn P., Brooks S., Buehler E., Chao Q., Khan S.,
 RA Kim C., Shinn P., Altafi H., Bel O., Chin C., Chou J., Choi E.,
 RA Conn L., Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B.,
 RA Lee J., Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N.,
 RA Nguyen M., Palm C., Pham P., Sakano H., Liu S., Southwick A.,
 RA Thaveri A., Tortum M., Vaysberg M., Yu G., Federspiel N.A.,
 RA Theologis A., Ecker J.R.:
 RT "Genomic sequence for Arabidopsis thaliana BAC T24P13 from chromosome
 1."
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Ecker J.R.:
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Ecker J.R.:
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Ecker J.R.:
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
 RA Khan S., Kim C., Altafi H., Bel B., Chin C., Chou J., Choi E.,
 RA Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,
 RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsky N.,
 RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
 RA Thaveri A., Tortum M., Vaysberg M., Yu G., Davis R., Federspiel N.,
 RA Theologis A., Ecker J.:
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Cho H., Cosgrove D.J.:
 RT "Expansin ALEXP10 affects organ growth and morphology in Arabidopsis
 thaliana."
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A.

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 11, 2002, 14:51:00 ; Search time 19.6871 Seconds
(without alignments)
2003.488 Million cell updates/sec

Title: US-09-896-301-2

Perfect score: 1259
Sequence: 1 AGCGWVNAHATFFGGDASG.....NNXNAGWSFGOTFGAHR 228

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database :

SPREMBL_19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1246	99.0	237	10	Q38863
2	1246	99.0	250	10	Q9C554
3	1139	90.5	249	10	Q9LDR9
4	1137	90.3	252	10	Q9FUM2
5	1120	89.0	249	10	Q93XP1
6	1115	88.6	248	10	Q80622
7	1106.5	87.9	249	10	Q9M515
8	1092	86.7	249	10	Q92P36
9	1090	86.6	247	10	Q9M517
10	1051.5	83.5	242	10	Q9LDB1
11	1049	83.3	250	10	Q9FVG9
12	1047	83.2	222	10	P93493
13	1047	83.2	253	10	Q9SMY1
14	1046	83.1	245	10	Q9LLB2
15	1042	82.8	232	10	P93492
16	1037	82.4	253	10	Q93XP2

17	1037	82.4	254	10	Q82093	082093 prunus arime
18	1037	82.4	260	10	Q9FNT0	Q9FNT0 cicer ariet
19	1036	82.3	232	10	P93495	P93495 pinus taeda
20	1034	82.1	254	10	Q9FUM3	Q9FUM3 prunus aviu
21	1029	81.7	249	10	Q9ZP35	Q9ZP35 nicotiana l
22	1028.5	81.7	252	10	Q9FS30	Q9FS30 prunus pers
23	1028	81.7	232	10	P93494	P93494 pinus taeda
24	1027	81.6	253	10	Q9SRT1	Q9SRT1 fragaria an
25	1026.5	81.5	252	10	Q81133	Q81133 prunus arime
26	1026	81.3	250	10	Q93625	Q93625 cucumis sat
27	1024	81.3	247	10	Q82625	Q82625 lycopersico
28	1020	81.0	276	10	Q94KT6	Q94KT6 zea mays (m
29	1010	80.2	239	10	Q9ZP31	Q9ZP31 lycopersico
30	1001	79.5	246	10	P93442	P93442 oryza sativ
31	1000	79.4	246	10	Q945J0	Q945J0 oryza sativ
32	996	79.1	291	10	Q9M4X8	Q9M4X8 oryza sativ
33	994	79.0	251	10	Q946J1	Q946J1 oryza sativ
34	993	78.9	253	10	Q94KT7	Q94KT7 zea mays (m
35	989	78.6	253	10	Q9SMD4	Q9SMD4 rumex palus
36	985	78.2	255	10	Q9FMA0	Q9FMA0 arabidopsis
37	979	77.8	251	10	Q40636	Q40636 oryza sativ
38	976.5	77.6	253	10	Q22874	Q22874 arabidopsis
39	973.5	77.3	245	10	Q9FNT1	Q9FNT1 cicer ariet
40	973	77.3	220	10	Q9AYR1	Q9AYR1 eustoma gra
41	968	76.9	258	10	Q49194	Q49194 gossypium h
42	955	75.9	257	10	Q48818	Q48818 arabidopsis
43	938	74.5	257	10	Q9SD24	Q9SD24 marsilea qu
44	936	74.3	257	10	Q9XG16	Q9XG16 lycopersico
45	931.5	74.0	252	10	Q9FY30	Q9FY30 festuca pra

ALIGNMENTS

RESULT 1
ID Q38863 PRELIMINARY: PRT: 237 AA.

AC Q38863: 01-NOV-1996 (TREMBL) 01. Created)
DT 01-NOV-1996 (TREMBL) 01. Last sequence update)
DT 01-DEC-2001 (TREMBL) 19. Last annotation update)
DE EXPANSIN AT-EXPI (FRAGMENT).
GN AT-EXPI.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=96016146; PubMed=7568110;
RA Shcherban T.Y., Shi J., Durachko D.M., Gultinan M.J.,
RA McQueen-Mason S.J., Shieh M., Cosgrove D.J.;
RT "Molecular cloning and sequence analysis of expansins--a highly
RT conserved, multigene family of proteins that mediate cell wall
RT extension in plants";
RL Proc. Natl. Acad. Sci. U.S.A. 92:9245-9249(1995).
DR EMBL: U30476; AAB38070.1; -;
DR InterPro: IPR000882; Pollen_allergen.
DR Pfam: PF01357; Pollen_allergen; 1.
DR PRINTS: PR01225; EXPANSINFAMLY.
DR PRODOM: PD002179; Pollen_allergen; 1.
FT NON_TER
SQ SEQUENCE 237 AA: 25155 MW: 055A2CB9C2943682 CRC64;

Query Match 99.0%; Score 1246; DB 10; Length 237;
Best local Similarity 98.7%; Pred. No. 1.5e-105;
Matches 225; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGCGWVNAHATFFGGDASGTGAGCGNLYSGYGTAAALSTALFPNNGLSGACAFET 60
DB 10 AGCGWVNAHATFFGGDASGTGAGCGNLYSGYGTAAALSTALFPNNGLSGACAFET 69